

OM protein - protein search, using sw model

Run on: May 28, 2002, 14:17:37 ; Search time 37.71 Seconds (without alignments)
 3550.731 Million cell updates/sec

Title: US-09-803-165-34

Perfect score: 4026

Sequence: 1 MIFDTDVYITKDOCKPPIRFK..... KEDLKVQSSKQVGLDAWLKK 774

Scoring table: BLOSUM62

Gapext 10.0 , Gapext 0.5

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0 %

Maximum Match 100 %

Getting first 45 summaries

Database : SPREMBL/19:*

1: sp_archea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_micr:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rabbit:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriapl:*

17: sp_archeapl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	3212	79.8	775	09P9K4 pyrococcus
2	3211	79.8	775	09HH06 Pyrococcus
3	3081	76.5	759	09HH98 BYQCGCCUS
4	1204.5	29.9	781	09P9N1 O9P9N1
5	1204.5	29.9	781	096YV1 096YV1
6	1190	29.6	803	059691 059691
7	1125	27.9	785	1 O9P9M2 O9P9M2
8	774	19.2	195	1 O9HH85 O9HH85
9	631.5	15.7	1443	5 O9NAHL O9NAHL
10	626	15.5	901	17 O9HRV9 O9HRV9
11	623	15.5	796	17 O9HJRO O9HJRO
12	622.5	15.5	800	17 Q97AH3 Q97AH3
13	616.5	15.3	876	17 Q971C7 Q971C7
14	616.5	15.3	1105	11 Q91vt0 Q91vt0
15	616	15.3	1107	4 Q96H98 Q96H98
16	584.5	14.5	914	1 Q959690 Q959690

17 567 14.1 1492 10 Q9F1A3 Q9f1A3 arabiopsis

18 559.5 13.9 223 17 026310 026310 methanother

19 555.5 13 1458 13 09d46 xenopus lae

20 527 13.1 1472 5 09vd90 drosophila

21 516.5 12.8 844 1 031096 cernarchaeum

22 509 12.6 845 1 074046 074046 cercarchaeum

23 507.5 12.6 1001 12 Q90626 Q90626 porcine lym

24 506.5 12.6 1004 12 Q90628 Q90628 porcine lym

25 506.5 12.6 1415 5 077034 077034 drosophila

26 504.5 12.5 2154 10 Q92VC8 Q92VC8 arabiopsis

27 491 12.2 913 12 Q9YZZ9 Q9YZZ9 chlorella v

28 491 12.2 1007 12 Q91F13 Q91F13 porcine cyt

29 488 12.1 2207 3 093845 093845 emericella

30 487 12.1 1007 12 Q91F15 Q91F15 porcine cyt

31 486 12.1 1007 12 Q91F16 Q91F16 porcine cyt

32 484 12.0 1855 5 Q9TX75 Q9TX75 plasmidum

33 484 12.0 1855 5 Q9BHNO Q9BHNO plasmidum

34 483 12.0 2271 10 Q9SGD5 Q9SGD5 arabiopsis

35 481 11.9 1007 12 Q91F17 Q91F17 porcine cyt

36 481 11.9 1009 12 Q9YTQ4 Q9YTQ4 arabiopsis

37 481 11.9 1912 5 Q9U0H1 Q9U0H1 plasmidum

38 479 11.9 1016 12 Q993K6 Q993K6 cellitrichi

39 478.5 11.9 1013 12 092827 092827 human herpe

40 477.5 11.9 774 2 Q9F175 Q9F175 pseudomonas

41 477 11.8 2243 4 Q9UNF3 Q9UNF3 homo sapien

42 477 11.8 2286 4 Q9Y554 Q9Y554 homo sapien

43 477 11.8 2297 4 Q9Y555 Q9Y555 homo sapien

44 471.5 11.7 1041 12 Q9DRT8 Q9DRT8 elephant he

45 468.5 11.6 1012 12 040910 040910 kaposi's sa

SEQUENCE FROM N.A.

RC STRAIN=AL585;

RC Querellou J., Cambon M.A., Lesongeur F., Forterre P., Barbier G.;

RA RT Thermococcales and phylogenetic implications";

RT Submitted (MAR 1998) to the EMBL/GenBank/DBJ databases.

RL - I - CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE - N

CC PYROPHOSPHATE + DNA(N).

CC - I - SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.

DR EMBL; AJ004834; CAB1809.1; -.

DR HSSP; P56589; 1TGO.

DR IntePrP; IPR02064; DNA_POL_B.

DR Pfam; PF00136; DNA_POL_B.

DR Pfam; PF03104; DNA_POL_B_exo; 1.

DR PRINTS; PR00106; DNAPOL_B.

DR SMART; SW00486; POLB; 1.

DR PROSITE; PS00116; DNA_POLYMERASE_B; UNKNOWN¹.

KW Nucleotidyltransferase; Transferase.

SEQUENCE 775 AA; 90492 MW; 677264920D7700F CRC64;

Query Match Similarity 79.8%; Score 3212; DB 1; Length 775; Best Local Similarity 77.1%; Pred. No. 1..le-180; Matches 595; Conservative 86; Mismatches 89; Indels 2; Gaps 1;

	1	-1	CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE - N PYROPHOSPHATE + DNA(N)
Db	1	MILDADYITEDGKPIIRIFKKEENGKFEVYDRNRPYIYALKKDDSAIDEKAIAKGRRH	60
Db	1	MILDADYITEDGKPIIRIFKKEENGKFEVYDRNRPYIYALKKDDSAIDEKAIAKGRRH	60
Qy	61	KIVRVDVAKKKFLGRPIEWKLYFEEHPODVAIRPAIDYDIFPAKRY	120
Db	61	KIVRIVDVKKKFLGRPIEWKLYFEEHPODVAIRPAIDYDIFPAKRY	120
Qy	121	LIDKGKIPMEGDEELKLLAFDIDTYLHGEEEAKGPIMISADEGAKVITWKVLDLY	180
Db	121	LIDKGKIPMEGDEELKLLAFDIDTYLHGEEEAKGPIMISADEGAKVITWKVLDLY	180
Qy	181	VDVSNEREMIRKFVQVREKPDVLITYNGNFDLILYLIKRAEKGTYLLGRDKHE	240
Db	181	VDVSNEREMIRKFVQVREKPDVLITYNGNFDLILYLIKRAEKGTYLLGRDKHE	240
Qy	181	VEVSSEREMIRKFVQVREKPDVLITYNGNFDLILYLIKRAEKGTYLLGRDKHE	238
Qy	241	PKTHRMGDSFAVEIKGRTHDIFPVWRITINLPTYLEAVVAVLGKTKSKLGAEETAAI	300
Db	241	PKTHRMGDSFAVEIKGRTHDIFPVWRITINLPTYLEAVVAVLGKTKSKLGAEETAAI	300
Qy	239	PKMQRGDMTAEIKGRHFDLYHVRITINLPTYLEAVVAVLGKTKSKLGAEETAAI	298
Qy	301	WETEESMKKLAQYSMEDARATELGKPFPMELAKLIGOSWMDVRSRSTGNLVEVYL	360
Db	299	WETGKGKLERVAKYSMEDAKVYELGRSPFPMQSLRSVLQGQLWDVRSRSTGNLVEVYL	358
Qy	361	RVAYERNELAPNPKDEEYRRRLRTTLYGGYKEPERGLWENITYLDFRCLPLSIWTH	420
Db	359	RVAYERNELAPNPKDEEYRRRLRTTLYGGYKEPERGLWENITYLDFRCLPLSIWTH	418
Qy	421	VSPDTLREGCKNYDAPVYKFCDFPGFPLSLLRKLDRDQEIKRMRKMAKDPIEK	480
Db	419	VSPDTLREGCKNYDAPVYKFCDFPGFPLSLLRKLDRDQEIKRMRKMAKDPIEK	478
Qy	481	MLDYQRRAVLHANSYYGGMCPKARWYKCAESVTAWGRHYIEMTIKEEKFCKV	540
Db	479	MLDYQRRAVLHANSYYGGMCPKARWYKCAESVTAWGRHYIEMTIKEEKFCKV	538
Qy	541	YADTGFATIGKEPEMKKAKFELYINSKPLGILELEVYEGFYRGRFFVAKKAVI	600
Db	539	YADTGFATIGKEPEMKKAKFELYINSKPLGILELEVYEGFYRGRFFVAKKAVI	598
Qy	601	DEGRTTRGLEYVRDWSLAKETQVKLEALKDSEVKAIVEIVKDVVEIYAKQVPL	660
Db	599	DEGKLTTRGLEYVRDWSLAKETQVKLEALKDSEVKAIVEIVKDVVEIYAKQVPL	658
Qy	661	EKLVHQITQKOLSEYKAIGPHVAIAKRLAAGKIKVPRGTISIYIVLGRSGKLSDRVLL	720
Db	659	EKLVHQITQKOLSEYKAIGPHVAIAKRLAAGKIKVPRGTISIYIVLGRSGKLSDRVLL	718
Qy	721	SEYDPKKHHYDDYIENQVLPAVRLAFCYRKELKLYQSKQVGDawl	772
Db	719	SEYDPKKHHYDDYIENQVLPAVRLAFCYRKELKLYQSKQVGDawl	770
RESULT	2		
QPHH06			
ID	QPHH06	PRELIMINARY;	PRM;
AC	QPHH06;		775 AA.
DT	01-MAR-2001 (TREMBLrel. 16, Created)		
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	DNA POLYMERASE.		
GN			
GN			
OS	Pyrococcus glycovorans.		
OC	Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.		
OX	Querelle, J. J. E., Cambon, M. A., Lessongeur, F., Barbier, G.;		
NCBI_TAXID=4610;			
RN	[1] SEQUENCE FROM N. A.		
RP			
RC	STRAIN=AL646;		
RC			
RC			
Qy	541	YADTGFATIGKEPEMKKAKFELYINSKPLGILELEVYEGFYRGRFFVAKKAVI	600
Db	539	YADTGFATIGKEPEMKKAKFELYINSKPLGILELEVYEGFYRGRFFVAKKAVI	598
Qy	601	DEGRTTRGLEYVRDWSLAKETQVKLEALKDSEVKAIVEIVKDVVEIYAKQVPL	660
Db	599	DEGKLTTRGLEYVRDWSLAKETQVKLEALKDSEVKAIVEIVKDVVEIYAKQVPL	658
Qy	661	EKLVHQITQKOLSEYKAIGPHVAIAKRLAAGKIKVPRGTISIYIVLGRSGKLSDRVLL	720
Db	659	EKLVHQITQKOLSEYKAIGPHVAIAKRLAAGKIKVPRGTISIYIVLGRSGKLSDRVLL	718
Qy	721	SEYDPKKHHYDDYIENQVLPAVRLAFCYRKELKLYQSKQVGDawl	772
Db	719	SEYDPKKHHYDDYIENQVLPAVRLAFCYRKELKLYQSKQVGDawl	770
RESULT	3		

PS Disclosure: Fig 9; 40pp; English.

XX
XX
CC
CC
CC
CC
CC
CC
CC
XX
SQ

The present invention relates to thermostable mutant B-type DNA polymerases, which have a Y-GG/A amino acid motif between the N-terminal 3'-5' exonuclease domain and the C-terminal polymerase domain, where the tyrosine of this motif is mutated. The mutant B-type DNA polymerase is useful for synthesising nucleic acids and for PCR. The present sequence is the protein sequence for a recombinant Thermococcus aggregans (Tag) DNA polymerase, which was used to illustrate the invention.

Sequence 774 AA;

Query Match 100.0%; Score 4026; DB 22; Length 774;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 774; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MIFDTDTKTKGKPKTRIFKENGKIELDPHFOQYIYALLKKDSADETAKAIGERHG 60
1 midtctyitkdgkpkirifkengkieleldphfqiyallkkdsadeikalkgerhg 60
61 kivrvvdavkvvkfkfgrdavewkvlifhpdavpkrlgkirehpavidiyeypdifikry 120

61 LIDKGKIPMEGDEELKLMARDFETFYHEGDFGKGIMISYADEEARITWKNDPY 180
1 lidkgkimpmegdeelklimafdietylhegdfgkgeimisyadeearitwkndpy 180

Db 181 VDVVSNSREMKRKFQVIREKDPDVILTYNGDNFDLPIYLKRAEKLGVTLLGROKEHPE 240
181 vdvsnsremkrkfqivrekdpdviltingndfdlylkrakekgvtllgrkhep 240

OY 241 PKIHRMGDSFAVEIKRGRHDFLPVVRRTILPITYLEAVVEAVIGKTSKLGAESTAAI 300
241 pkhrmgdsfaeveikrgrhdflpvrrtinpityleavveavigktksklgaaiai 300

Db 301 WETEESMKKLAQYSMEDARATYELGKESKEFPEMAELAKLIGSWDVSRSSTGNLWYLL 360
301 weteesmkklaqysmedaratyelgkeffpmeaelaklqgswdvsrsstgnlwyl 360

OY 361 RVAYERNELANPKPDEEYRRRLRTYLGGYKEPERGLMENTYDFRCIYPSITHN 420
361 rvayernelanpkpdeeyrrrlrtylggkykeperglmentydfrciypsithn 420

Db 421 VSPDTLREGCKNYDAPIVGKFCDFPGFIPSTIGELTMRQETKKMKTIDPIKK 480
421 vspdtleregcknydapivgkfcdfpgfipstigeltmrqetkkmkatidpikk 480

OY 481 MLDYRGRAVKLUHANSYGYMGYPKARWYSKCAESTAWGRHYIEMTIKEEKGFKVL 540
481 mldyrgqravkhansygymgypkarwyskcaestawgrhyiemtikeekgfkvl 540

Db 541 YADTGTYATIGKEPEKTIKKAKERLYKINSKPLLELEYEFGYLGFVAKKRYAVI 600
541 yadtgtyatigkepektikkakelerlykinskplleleyefgylgfvakkravyi 600

Db 601 DEEGRITTRGLEVVRDWESETAKETGKVLAILKEDSVEKAETKWDVWEAKVQPL 660
601 deegrirtrglevvrdwesetaketgkvlailkedsvekavelvkvdeeyqypl 660

OY 661 EKLVHQITKDLSEKAIGPHVIAKRLAAGKIKYRPGTISYYTURGSKISDRVILL 720
661 eklvhqitkdlsekaigphvialakrlaagkikyrgtisyyturgskisdrvill 720

Db 721 SEYDPPKHKYPPYYENQVPAVRLTEAIGYRKEDKQYKQSGQVGDWALK 774
721 seydkhkypdyyenqvipavrlileafgyrkedkqyqsgqvgdawlk 774

XX
XX
AC AAW29323;
XX
XX
DE 20-APR-1998 (first entry)
XX
DE DNA polymerase with 3'-5' exonuclease activity.
XX
KW TY Exon; DSM 10597; thermostable; DNA polymerase;
3'-5' exonuclease; amplification.
XX
OS Thermococcus sp.
XX
RN DE19611759-A1.
XX
PD 02-0CM-1997.
XX
PR 25-MAR-1996; 96DE-1011759.
XX
PA (BOEFL) BOHRINGER MANNHEIM GMBH.
XX
PI Antranikian G, Frey B, Niehaus F;
XX
DR WPI; 1997-481494/45.
XX
DR N-PSDB; AAT86434.
XX
PT Thermostable DNA polymerase from Thermococcus sp. TY - useful for
PT nucleic acid amplification
XX
PS Claim 1; Pages 9-10; 32PP; German.
XX
CC The present sequence (TY Exon) is a Thermococcus sp. TY - useful for
CC (DSM 10597) thermostable DNA polymerase with 3'-5' exonuclease
CC activity.
CC The enzyme can specifically amplify nucleic acid fragments of up to
CC 5 Kb in high yields, has an activity half-life of 20 minutes at 90
CC degrees C, has an optimum temperature of 70-80 degrees C, has an
CC optimum pH of 7.5, exhibits optimum activity at a KCl concentration
CC of 80-100 mM, is magnesium ion-dependent and is inhibited by
CC manganese ions.
XX
SQ Sequence 774 AA;

Query Match 99.1%; Score 3990; DB 18; Length 774;

Best Local Similarity 99.4%; Pred. No. 0; Mismatches 1; Indels 4; Gaps 0; Gaps 0;

Matches 769; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 MIFDTDTKTKGKPKTRIFKENGKIELDPHFOQYIYALLKKDSADETAKAIGERHG 60
1 midtctyitkdgkpkirifkengkieleldphfqiyallkkdsadeikalkgerhg 60
61 kivrvvdavkvvkfkfgrdavewkvlifhpdavpkrlgkirehpavidiyeypdifikry 120

61 LIDKGKIPMEGDEELKLMARDFETFYHEGDFGKGIMISYADEEARITWKNDPY 180
1 lidkgkimpmegdeelklimafdietylhegdfgkgeimisyadeearitwkndpy 180

Db 121 VDVVSNSREMKRKFQVIREKDPDVILTYNGDNFDLPIYLKRAEKLGVTLLGROKEHPE 240
121 vdvsnsremkrkfqivrekdpdviltingndfdlylkrakekgvtllgrkhep 240

OY 241 PKIHRMGDSFAVEIKRGRHDFLPVVRRTILPITYLEAVVEAVIGKTSKLGAESTAAI 300
241 pkhrmgdsfaeveikrgrhdflpvrrtinpityleavveavigktksklgaaiai 300

Db 241 pkhrmgdsfaeveikrgrhdflpvrrtinpityleavveavigktksklgaaiai 300

OY 301 WETEESMKKLAQYSMEDARATYELGKESKEFPEMAELAKLIGSWDVSRSSTGNLWYLL 360
301 weteesmkklaqysmedaratyelgkeffpmeaelaklqgswdvsrsstgnlwyl 360



RESULT 2

AAW29323

AAW29323 standard; Protein; 774 AA.

GenCore version 4.5
copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 28, 2002, 14:17:02 ; Search time 34.46 Seconds
(without alignments)
2494.809 Million cell updates/sec

Title: US-09-803-165-34

Perfect score: 4056

Sequence: 1 MIFDTBYITKDGKPIRIFK.....KEDLKYQSSKOVGLDAWLKK 774

Scoring table: BLOSUM62

Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Actual number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802:*

1: /SRS1/gcdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*

2: /SRS1/gcdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*

3: /SRS1/gcdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*

4: /SRS1/gcdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*

5: /SRS1/gcdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*

6: /SRS1/gcdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*

7: /SRS1/gcdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*

8: /SRS1/gcdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*

9: /SRS1/gcdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*

10: /SRS1/gcdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*

11: /SRS1/gcdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*

12: /SRS1/gcdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*

13: /SRS1/gcdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*

14: /SRS1/gcdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*

15: /SRS1/gcdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*

16: /SRS1/gcdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*

17: /SRS1/gcdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*

18: /SRS1/gcdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*

19: /SRS1/gcdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*

20: /SRS1/gcdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*

21: /SRS1/gcdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*

22: /SRS1/gcdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	4026	100.0	774 22 AAG78941	Recombinant Tag DNA polymerase with T. litoralis DNA polymerase motif
2	3990	99.1	774 18 AAW29323	Thermococcus gorgo
3	3446.5	85.6	1829 18 AAW29322	Thermococcus gorgo
4	3356	83.4	1702 14 AAR8095	Thermococcus gorgo
5	3232.5	80.3	773 19 AAW06285	Thermococcus gorgo
6	3231.5	80.3	773 19 AAW8818	Thermococcus gorgo
7	3226.5	80.1	774 18 AAW3111	Thermococcus pepto
8	3225.5	80.1	774 19 AAW11313	Heat-resistant DNA
9	3222.5	80.0	774 17 AAR3154	Mature DNA polymerase
10	3222.5	80.0	774 18 AAW0048	KOD thermostable
11	3221.5	80.0	774 19 AAW48453	Mutant KOD DNA pol

RESULT 1

ID AAG78941 standard; Protein: 774 AA.

AC AAG78941; XX

DT 13-FEB-2002 (first entry)

DE Recombinant Tag DNA polymerase.

XX Tag polymerase; B-type DNA polymerase.

XX OS Thaumococcus aggregans.

XX PN EP1132474-A.

XX PD 19-SEP-2001.

XX PR 06-MAR-2001; 2001EP-0104583.

XX PR 11-MAR-2000; 2000EP-0105155.

PA (HOFF) ROCHE DIAGNOSTICS GMBH.

PA (XX) Sobeck H, Frey B, Antranikian G, Boehlke K, Pisani FM, Rossi M; XX WPI; 2001-618367/72.

DR DR N-PSDB; AAI70053.

XX PT New thermostable mutant B-type DNA polymerase with a Y-GG/A amino acid motif between the N-terminal 3'-5' exonuclease domain and the C-terminal polymerase domain in the wild type polymerase, useful for synthesizing nucleic acids

QY 361 RVAVERNELAPNKDEEYRRRLTYLGGYKPERGLMENTYDFCILPSIVTHN 420
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 361 rvayernerelapnkdeeyrrrltylggykperglmentydfcilspsivthn 420
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Qy 421 VSPDTLEREGCKKNVYDAPVYGVKPFKDFGFPSTIGELTMROPIKKMATAPIRK 480
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 421 vspdtleregcknydavpvygvkpfkdfgfpstigeltmropiikkmatdpielk 480
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Qy 481 MLDY'KORAVLHANSYGYMGKPYKARWYKRECAESVTANGRHYIEMTIEERFGFYL 540
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 481 mldy'qrawllansygymgkpykarwyskecaesvtawgrhyiemi:keleekfkgkv1 540
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Qy 541 YADPDGFYAVIPEKPETRKKAKREFLKYINSKPGLELEYEGFYLRFPPVAKRYAVI 600
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 541 yadtogfyatipgekpetikkakeflykynskpgleleyedfyiqlfvaakryavi 600
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Qy 601 DEEGITTRCLEVVRDWSIAKEQAKVLEALKEDSEKAVETVKDVEERAKYQVL 660
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 601 deegittrclevvrdsweiakeqakvlealkedsekaevetvkdvveerakyqvl 660
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Qy 661 EKLVTHEQTKDLSYFKAIGPHVATKRLAAKGKVKRPGTISIVLRLSGKSDRVL 720
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 661 ekvlvheqtkdlsyfkaigphvatakrakgkvrpgtisivlrlsgksdrvl 720
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Qy 721 SEYDPEKKHIDPDYTIENQYLPAYLRLIAFGYKREDIYQSYKQVGLPAWKK 774
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 721 seydpkkhidyptienqylpaylrliafgykrediyqsykqvglpaawkk 774
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 RESULT 3
 ID AAW29322 standard; Protein; 1829 AA.
 XX AAW29322;
 AC
 DT 20-APR-1998 (first entry)
 DE DNA polymerase with 3'-5' exonuclease activity.
 XX TYPOL Intron; DSM 10597; thermostable; DNA polymerase;
 KW 3'-5' exonuclease; amplification.
 XX Thermococcus sp.
 OS
 FH Key Location/Qualifiers
 FT Misc-difference 116
 FT /label= "unknown"
 FT /note= "encoded by GGN"
 FT 118
 FT /label= "unknown"
 FT /note= "encoded by NTC"
 FT 1123
 FT /label= "unknown"
 FT /note= "encoded by NTC"
 XX DEL19611759-A1.
 XX 02-OCT-1997.
 PN 25-MAR-1996; 96DE-1011759.
 PR
 PD 25-MAR-1996; 96DE-1011759.
 PA
 PT Antanikian, G., Frey, B., Niehaus, F.;
 DR WPI; 1997-481494/51
 DR N-PSDB; AAT86433
 PT Thermostable DNA polymerase from Thermococcus sp. TY - useful for
 PT nucleic acid amplification
 XX

PS Claim 1; Pages 5-8; 32pp; German.
 XX
 CC The present sequence (TYPOL Intron) is a Thermococcus sp. TY (DSM 10597) thermostable DNA polymerase with 3'-5' exonuclease activity.
 CC The enzyme can specifically amplify nucleic acid fragments of up to 5 kb in high yields, has an activity half-life of 20 minutes at 90 degrees C, has an optimum temperature of 70-80 degrees C, has an optimum pH of 7.5, exhibits optimum activity at a KCl concentration of 80-100 mM, is magnesium ion-dependent and is inhibited by manganese ions.

SO Sequence 1829 AA;

Query Match 85.6%; Score 3445.5; DB 18; Length 1829;
 Best Local Similarity 42.8%; Pred. No. 4.7e-260; Mismatches 0; Gaps 3; Indels 1055; Gaps 3;

DB 1 MIFDDTYIKDGKPIRKFKENGFKEIDPHQPYTAWALKSDAIDKATKGERIG 60
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 1 mldy'qrawllansyfkaigphvatakrakgkvrpgtisivlrlsgksdrvl 720
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 61 KIVRTVDAVKKRFLGRDVEWKLFERHQDVALRKIREHRAVIDYEYDIPAKY 120
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 61 kivrvdavvkkflgrdvewklfenpqdvpalrkirehpaividyeypakry 120
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 121 LIDKELIPMEGDEEIKLMDIEPYHEDFGKELIMISADEEAVITWNIDLEY 180
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 121 lidkeliipmegdeelklimadiepyhedefgkeliimisydeeeavitwnidley 180
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 181 VDVVSNEREMIKRFOVIRVKDPDYLITNGDNEDPLVLIKRAEKLGYTLLGRDEKE 240
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 181 vdvvsneremikrfovirkdpdylitngdnedplvlikraeklgvtllgrdkepe 240
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 241 PKIHRGDDFAVELRGRIFDLEPVVRITNLPTYTLAEVYAVLGTKTSKLGEETIAI 300
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 241 pkihmgdafaevskgrindalpvriftnlpvtytleavylgktsklaeiai 300
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 301 WETEESMKIQLAQSNEDARATYELGKEFPMEAELAKLIGOSWDSSTGNUEWTL 360
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 301 weteesmkiklaqyamedatarelyelgkefpmeaelsaklgqswdsstgnuewtl 360
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 361 RVAVERNELAPNKDEEYRRRLTYLGGYKPERGLMENTYDFCILPSIVTHN 410
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 361 rvayernerelapnkdeeyrrrltylggykperglmentydfcilspsivthn 420
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 411 ----- 410
 Db 421 gkgivnisdvkeydylglgwgwqrvkvvkyhyegklininglctpnhkpvtendq 480
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 411 ----- 410
 Db 481 tridslakslfskskvgkliittkfekeafeknkpsseelkgelsgsilaegtlrk 540
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 411 ----- 410
 Db 541 dieyfdssrkkrishqyveitigenekellerilyifdklgirpskkdgtnalki 600
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 411 ----- 410
 Db 661 kildslgipysrykyiengkeltkhlleitgrdglfqtlglfslssekanalekiae 720
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 411 ----- 410
 Db 601 ttakkavylqieelknieslyapavrgrfferdatvnkirstiwtqtnnkwdiwa 660
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 411 ----- 410
 Db 721 vremrlnknsfylstfesseykgevydltlegnpyfangilthnslypsivthn 780
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 421 VSPDTLEREGCKKNVYDAPVYGVKCFKDFGFFISLGETIMTROEIKKMKATDPIRK 480
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 781 vspdtleregckhydavivgyifckdpfipspilgelitmrqeiikkmmkattidplekk 840
 Qy 481 MLDYRQRRAVAKLHA-.....
 |||||||
 Db 841 mldyqravkllanslipnewipilengevfkfkgifedrimeeqkdkvrtvdtevle 900
 Qy 494
 |||||||
 Db 901 vdnlafaslnkeskkseikkvakalrhkkygeayeevlnsgrkhiitrghsflftrngki 960
 Qy 494
 |||||||
 Db 961 kelwgeevkvqgdlivpkvklnekeavainpeliskipdedtadvvmttppkgrknffk 1020
 Qy 494
 |||||||
 Db 1021 gmlrtlkwifgeeskirktnrylfhlelgfkwkllprgevtdwgklyrqllyeklv 1080
 Qy 494
 |||||||
 Db 1141 vsgayagaqkntggmsysvkllynenpnvklomklaekffgkvrvgknvcvdkpkmay 1200
 Qy 494
 |||||||
 Db 1261 vfilnslgvssikigfsgvvyrynedlpflqtsrqntypnlpkevleefgrkfq 1320
 Qy 494
 |||||||
 Db 1321 knlfekkeladsgkdkrkkldflngdvlavknyekregeywydsvnedn 1380
 Qy 494
 |||||||
 Db 1381 fivgfglyahnsyggymgypkarwyskecaestvtawgrhylentlikeekfgfkwya 1440
 Qy 543 D-.....
 |||||||
 Db 1441 dsvtdgdelivkrngnriefvpieklfervdyrigekeyciledveatldnrgkliwky 1500
 Qy 544
 |||||||
 Db 1501 PymrhrakkyyriwtswyldvtdhsLivaedglkearpmelgekslatkdsg 1560
 Qy 544
 |||||||
 Db 1561 veykphaleeisyyngvvydlevgthrffanglvhntdfyatiipgekpetikkake 1620
 Qy 566 FLKYINRKLPGLELLEVYEGYRGEFAKKVAVIIEGRITTRGELEVRDWESETAKET 625
 |||||||
 Db 1621 fkyinrklpglelleyegyifrgfvakryavideegriftrglevvrdwseaket 1680
 Qy 626 QAKVLEALKESVEKAEVTKDQVVEBIAKQWPLKLVHREQIKDSEYKAIGHVAT 685
 |||||||
 Db 1681 gakvlealkesvekavetkdkdseyaikgphvai 1740
 Qy 686 AKRLAAGKIKVPGTISYIVRGSKISDRVLLSEYDPKKHKKYDPPYYTENQVPAVL 745
 |||||||
 Db 1741 akrlaaqkikvpgtisylivgskisdrvllseydpkkhkydpYYienqvlavl 1800
 Qy 746 RILEAFYRKEDLYKOSKQVGLDAMKK 774
 |||||||
 Db 1801 rileafyrgkqdkyqvgldawlik 1829
 RESULT 4
 AR38095

ID AAR38095 standard; Protein: 1702 AA.
 XX
 AC AAR38095;
 XX
 DE T. litoralis DNA polymerase.
 XX
 XX DNA polymerase; intron; translation; bacteriophage NEB #619; E. coli;
 KW mutation; cell growth; endonuclease; self splicing; stability;
 KW expression; transformation; expression vector.
 XX
 OS Thermococcus litoralis.
 XX
 PN EP547920-A.
 XX
 PD 23-JUN-1993.
 XX
 PF 18-DEC-1992; 92EP-0311622.
 XX
 PR 18-DEC-1991; 91US-0811421.
 XX
 PA (NEW) NEW ENGLAND BIOLABS INC.
 XX
 PI Comb DG, Jack WE, Kucera R, Perler F;
 XX
 DR WPI; 1993-198889/25.
 XX
 DR-NPSDB; AAQ43515.
 PR Recombinant thermostable DNA polymerase from archaeabacteria -
 expressed efficiently by removal of intervening introns
 XX
 PS Claim 1; Fig 19; 63pp; English.
 XX
 CC This sequence represents T. litoralis DNA polymerase. The DNA
 sequence encoding this protein contains introns described in the
 specification, however this protein sequence appears to be translated
 through these intronic sequences. Only a portion of the DNA
 polymerase protein sequence is reproduced in the specification. The
 T. litoralis DNA polymerase gene sequence may be derived from
 bacteriophage NIEB #619 on an approx. 14Kb BamHI restriction fragment.
 CC When this gene is transformed in to E. coli and expressed in its
 entirety, it is found to be unstable. There is a high frequency of
 mutation in the polymerase gene, cell growth is slow and there is some
 degree of cell mortality. This instability is due in part to the
 presence of the two introns. It is thought that the intervening DNA
 in this sequence encodes a separate protein which may be spliced out
 at the protein level. The second, 1170bp, intron has been found to
 code for an endonuclease which self splices out in E. coli. To
 improve stability of the DNA when it is expressed in E. coli, the
 specification states that it is desirable to delete the introns
 CC before transformation of the expression vector into E. coli.
 XX
 SQ Sequence 1702 AA;
 Query Match 83.4%; Score 3356; DB 14; Length 1702;
 Best Local Similarity 42.7%; Pred. No. 5e-253;
 Matches 726; Conservative 33; Mismatches 215; Indels 928; Gaps 2;
 Matches 726; Conservative 33; Mismatches 215; Indels 928; Gaps 2;
 Qy 1 MIFDTDYLTDGKPTIRIFKENGKFKIEDLPHCPYIALLKODSAJDEIKAKGERH 60
 |||||||
 Db 1 mldtayitkdgkplirifkengfkielphcpqiyallkodsaileeikakgerh 60
 Qy 61 KIVRVDVAKKKFLGRDVEWKLIFERHQDVALRGKREHPAVIDYEYDIPFAKKY 120
 |||||||
 Db 61 ktrvxlavwkrflgrevewkifbqdpamgrgslrebavvdyyeydpfaky 120
 Qy 121 LIDKGKIPMCGDEELKLMADIETYHEGEGFGEIIMIISYADBEEARVITWNKIDPY 180
 |||||||
 Db 121 lidkglipmgeedekilaadietyhegqefgkgeimisadyeearvitwnkidy 180
 Qy 181 VDWNSNERMKRFVQIVRKDPDVLYTGNDNEDPLVYLIKRAKLGVILLGRKERE 240

Db 181 vdvvsneremirkvfvqvkedpkdoviltynghdfplikrakqvrivlgdkehpe 240
 Qy 241 PKHRMGDSFAVEIKGRHFDFLPPVRITINPITYLEAVFALGKTKSKGAETRAI 300
 Db 241 Pkfqrmqdsfaveikgrhfdfpvrntinptyleavlgitskqgaaei 300
 Qy 301 WETEESMKKLAQSMEDARATFELGKEFPMEELAKLJGOSWWDVRSRSTGNLVEWLL 360
 Db 301 weesmkklacysmederatylegkfpmeelaklqgswdvsststnlyewll 360
 Qy 361 RVYERNELAPNHPDEEEYRRRTYIGYYKEPEKGWLWNTYDFRCLLFSTIYHN 420
 Db 361 rvayarnelapnkpdeeevkrrrrtyigyykepekgwlwntydfrcllf 420
 Qy 421 VSDTLEREGCKKWDVAPVGYKFCQPGFPPSILGELITRMRQEEIRKKMKTIDPEKK 480
 Db 421 vspdtlelegckydvapvgyrkfcktpfpgfipsgdiamqdkmmstidplekk 480
 Qy 481 MLDYRQRAVKLHA- 493
 Db 481 mldyqrakilanslipnewpliengelkfkigerfinsymekqenkvtenteyle 540
 Qy 494 ----- 493
 Db 541 vnnlfafsnkkikesevkkvalirkkygkayeqlssgrkinitaghsfltvngel 600
 Qy 494 ----- 493
 Db 601 kevsgdgikgedglivapkklklnkqgvininpelisdlseeetadimtisakgrkuff 660
 Qy 494 ----- 493
 Db 661 gmrlrtlrwmfgeenrrirtfnrylfhlekgliklprgyevtdwerlkkyqlyeklag 720
 Qy 494 ----- 493
 Db 721 svkyngnreylvmfneikdfisypqkeleewkgigtingfrnc1kvdedffgkllggy 780
 Qy 494 ----- 493
 Db 781 vsgyagaqknktggisysvkvynedpvnlesmkvnvaekffkkvrdrvncsiskkayl 840
 Qy 494 ----- 493
 Db 841 vnkclcgalaenkripsvilstspprvwsfleayftgddihpskrflstksellanol 900
 Qy 494 ----- 493
 Db 901 vflnslgissvkgfsgvyyrvinelqfpqtsrekntyysslpeiklrdvgkfq 960
 Qy 494 ----- 493
 Db 961 kmntffkkelvdsqklnrekakkleffingdivldvksvkekdyyggyvalsvnedn 1020
 Qy 494 -----NSYGYNGYPKRWYSEKAESVTANGRHYIEMTKEIEKEFGFKVLYA 542
 Db 1021 fvgfgllyahnsyggymgypkarwyskceasvtaawgrhjentireeekfgfkviya 1080
 Qy 543 D----- 543
 Db 1081 dvsgeseiirrqngirkvikkdkfsvdysigekyclevealtlddgk1vwkp 1140
 Qy 544 ----- 543
 Db 1141 pyvvrhrankrurifwltwsvidteds19ylntsktakkgerlkvwpfegk 1200
 Qy 544 ----- 543
 Ps ----- 543
 Db 1201 avkslicpnaplkdentktselavkfvelvgivingdgnwgdsraeyylqstgkdaee 1260
 Qy 544 ----- 543

Db 1261 ikqkllpleptkydvisnyyppknekdfnlakslvfmkrhfkdegrkrkipefmylpv 1320
 Qy 544 ----- 543
 Db 1321 tyleafrlgifsadgtvtirkgypeirltnidafirevrklwliwqinsifaettpr 1380
 Qy 544 ----- 543
 Db 1381 yngvstgtykskhirknkwrfraerigflerkqrlehlksarvkrntidfgfdlyhv 1440
 Qy 544 ----- 543
 Db 1441 kveeipyegyydileveethrrffannilvhntdgfyatipgekpelikkakefnyins 1500
 Qy 573 KLGGLLELEYEGYLGSFFAKARYAVDEDEGRITRDLVEVRDWSSTAKEMQAKLEA 632
 Db 1501 kipgileleyegyrlgfvtkkryavideegriftrtglevrvrdwseiaketqakvlea 1560
 Qy 633 ILKEDSVKAVELVKDVEIAYKQVPLKUHETOITKDLSEYKATGPVHATKRAAK 692
 Db 1561 ilkegsyekavevrdvwekaytrvpleklvheqitrdkkyqgphvaiakrtaar 1620
 Qy 693 GIKVRPGITISIVLRESGKISDRVILSEYDPRKKVHDYIENQVLPAVRILEAFG 752
 Db 1621 gikvpgtislyivkqgkisdrviliteydrkrhjydpdyiengvlpavrlileafg 1680
 Qy 753 YRKEDELIKTOSSKGVGLDWLKK 774
 Db 1681 yrkedlryqskqtgdawlk 1702

RESULT 5
 AA46285 ID AA46285 standard; Protein: 773 AA.
 XX
 AC AA46285;
 XX
 DT 28-SEP-1998 (first entry)
 XX
 DE Thermococcus gorganarius (Tgo) DNA polymerase protein.
 XX
 KW Thermostable; DNA polymerase; Thermococcus gorganarius; Tgo; screening;
 KW 3'-5' proofreading exonuclease; Taq polymerase; PCR; amplification;
 KW intervening sequence.
 XX
 OS Thermococcus gorganarius.
 XX
 PN EP034570-A1.
 XX
 PD 08-APR-1998.
 XX
 PF 03-OCT-1996; 96EP-0115874.
 XX
 PR 03-OCT-1996; 96EP-0115874.
 XX
 PA (BOEFL) BOEHRINGER MANNHEIM GMBH.
 XX
 PI Angerer B, Ankenbauer W, Bonch-osmolovskaya E, Ebenbichler C;
 PI Laue F, Schmitz-aghegian G, Svetlichny V;
 XX
 DR WPI; 1098-105468/18.
 DR N-PSDB; AAV26371.
 XX
 PT New isolated thermostable DNA polymerase - obtained from
 PT Thermococcus gorganarius, used for amplifying DNA or for DNA
 PT cloning, sequencing or labelling
 XX
 PS Disclosure; Page 16-21; 41pp; English.
 XX
 CC This amino acid sequence produces the thermostable DNA polymerase in
 CC Thermococcus gorganarius (Tgo). The DNA polymerase enzyme was isolated
 CC from this species by standard isolation and purification techniques. The
 CC thermostable enzyme possesses a 3'-5' proofreading exonuclease function.

this activity is not present in *Taq* polymerase which is normally used in PCR amplification. Thus this activity enables the fidelity of PCR to be increased, and therefore would no longer be prone to base incorporation. The *Tgo* DNA polymerase exhibits a two fold greater replication fidelity than any known polymerases, it can also decrease non-specific background amplification in PCR, due to its 3'-5' activity. The gene from this species also does not contain intervening sequences, which would have to be removed to enable expression in *E. coli* to occur. This enzyme has an approx molecular weight of 92.96 kb, and has been shown to retain 90% of its activity after incubation. In the presence of a stabilizer for two hours at 95 deg. C. This *Tgo* DNA polymerase can be used for e.g. amplifying DNA or for DNA cloning, sequencing or labelling.

XX
SQ Sequence 773 AA;

ULT 6
58818
AAW58818 standard; Protein: 773 AA.
AAW58818;
17-AUG-1998 (first entry)
Thermococcus gorgonarius thermostable DNA polymerase.
DNA polymerase; thermostable enzyme; DNA amplification; PCR;
cloning; sequencing; DNA labelling.
Thermococcus gorgonarius.

Query Match 80.3%; Score 3232.5; -DB 19; Length 773;
 Best Local Similarity 77.7%; Pred. No. 7.2e-244;
 Matches 601; Conservative 83; Mismatches 86; Indels 3; Gaps 2

08-APR-1998.
16-JAN-1997; 97EP-0100584.

03-OCT-1996; 96EP-0115874.
(BOEF) BOEHRINGER MANNHEIM GMBH.
Angerer B, Ankenbauer W, Bonch-Osmolovskaya E, Ebenbichler RC;

61 ttrvrvraekvkkfgrpiewklyfhpgqvpairdkikhenpavvdiyediptkry 1200
121 LDKGLIPMGEGBELKJMAFDITFYHGDERGKGETIMISADEEAVTWKNIDPY 1800
121 ldkgj1pmegoeeklmaldfetlynegeeaegpilmisradeegarvitwknidpy 1800
181 VVWVNSNREMRIRFVOTREKUPDVLTLYNGNFNDLYLIKAEKLTGVTLLGRDKHPE 2400
181 ||||| :|||||:::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 volvstekemirflikvkekekpdvillynqndnfdylfkksekiqkfqfgr-egse 2388

Laue F, Schmitz-Aghequian G, Svetlichny V; WPI; 1998-195469/18. N-PSDB; AAV11396. New isolated thermostable DNA polymerase - obtained from *Thermococcus gorgonarius*, used for amplifying DNA or for DNA cloning, sequencing or labelling.

This polypeptide comprises a novel thermostable DNA polymerase (TDP) of *Thermococcus gorgonarius*, a thermophilic archaeobacterium isolated from a thermal vent. The TDP catalyses the directed polymerisation of DNA and retains about 90% of its activity after incubation for 2 hr at about 95 deg C in the presence of a stabilizer. It exhibits more than 2-fold greater replication fidelity than DNA polymerase from *Pyrococcus furiosus*, and also has 3'-5' exonuclease (proofreading) activity. This exonuclease activity can decrease non-specific background amplification in PCR by degrading off-ends of primers bound to unspecific sequences. The TDP is suitable for use in DNA amplification, cloning, sequencing and labelling. The native enzyme can be used or a recombinant enzyme that has been expressed in microbial host cells transformed with a vector carrying an isolated DNA sequence (see AAV11396) coding for TDP.

Sequence 773 AA;

Db	181 vdvvstekemirkrkvkvedkdapdviilyngdnfotaylkkrkselgkvfillr--egse	CC	The present sequence is a Thermococcus peptonophilus 06-1
Qy	241 PKIHMGSFAVEIKGRIHDLFPRVRTRINPLTYLEAVLGTKSKLGAEEIAI	CC	(JCM 9653) derived thermostable DNA Polymerase, which can be used
Db	239 pkqmgdraveviqrihfdlypvrirntlpyleavtayeaifgkpekyeeiaa	XX	for nucleic acid sequence amplification, e.g. PCR.
Qy	301 WETESMKKLAQYSMEDARATYELGKEFFPMEAEELAKLIGQSVMDSRSSTGNLVEWL	XX	Sequence 774 AA;
Db	299 wetggelgrarvysmedaktyelgkfmpeaqslsrlvqslwvsrstgnlviewfl	XX	
Qy	361 RVAYERNEALAPNPKDDEEYRRRLRTYLYGGYVKEPERGLIMENTITYLDFCLYPSITVHN	420	
Db	359 rkayernelapnkpoderelarr-tesyagvykebergiweniyldfslypsiithn	417	
Qy	421 VSPDLEREGCKNYVAPTYGKCKDFGFGFIPSTIGELITMRQBIKKMKAIDPIEK	480	
Db	418 vspdlnredecceyavapvgkckdfgfgfipstigelitmrqbiikkmkaidpiek	477	
Qy	481 MDYQRAVKLHANSYYGKPYKARWYSKECAESVTAMGRHYIEMTIKEERGFKL	540	
Db	478 lldyqraikilansfyyggykarwykrecasvtagreylettireeefqkv1	537	
Qy	541 YADTGFATIPGEKPETIKKAKFELKYTKNSKPLGLELEYEGYLRLGFVAKRYAVI	600	
Db	538 yadtqffatipgadetvkkakfelynkpgaleleyegfykqffvttkkayvi	597	
Qy	601 DEESEITTGLEVVRDWBIAKEQAKYEALEIKAIDSEKAVEVLDVVEETKQVPL	660	
Db	598 deegitttglevvrdrwbiakeqarvleakgdkqfveavrikteklkyevpp	657	
Qy	661 EKLYTHEQTKDLSEYKAIGPHVATKRAAKGKIVRPGTISIVLRSGKLSDRVLL	720	
Db	658 eklyheqitrdksdykatgphavakriaargkirkptvsiyvlgksgrldraipf	717	
Qy	721 SEYDPRKKHIDPDYIENQVLPAVRILEAFGYRKEDLKQSSKQVGLDAWLK	773	
Db	718 defdpkhydaeyiengylpaverilrafgyrkdrqktrqvgisawlk	770	
RESULT	7		
ID	AAW33111		
AAW33111 standard; Protein; 774 AA.			
AC	AAW33111;		
XX			
DT	28-JAN-1998 (first entry)		
XX			
DE	Thermococcus peptonophilus OG-1 thermostable DNA polymerase.		
KW	JCM 9653; thermostable; DNA Polymerase;		
KW	nucleic acid sequence amplification; PCR.		
OS	Thermococcus peptonophilus.		
XX			
PN	JP09252776-A.		
XX			
PD	30-SEP-1997.		
XX			
PF	19-MAR-1996; 96JP-0063112.		
XX			
PR	19-MAR-1996; 96JP-0063112.		
XX			
PA	(TOY) TOYOB0 KK.		
XX			
WPI:	1997-530149/49.		
DR	N-PSDB; KAT88373.		
XX			
PT	Thermococcus peptonophilus thermostable DNA polymerase - useful for		
PT	nucleic acid sequence amplification, e.g. polymerase chain reaction		
XX			
PS	Claim 9; Pages 9-12; 27pp; Japanese.		
RESULT	8		
ID	AAW1313		
AAW1313 standard; Protein; 774 AA.			
AC	AAW1313;		
XX			
DT	20-MAY-1998 (first entry)		
XX			
DE	Heat-resistant DNA polymerase.		

Qy 481 MLDYRQRAVKLUHANSYGMGPKARWYKSCAESVTAWGRHYIEMTIKEEKGCFKL 540
 :|||||:||:|||||:||:|||||:||:|||||:||:|||||:||:|||||:||:|||||:||:|||||:||:||:
 Db 478 lldyraqraikllansyggryarwyckeasvytawgryeitntikeekygfki 537
 Qy 541 YADTGGYATPGKEPKETIKKKAKERFLKYNSKPLGPLELEYEGFLRGFVAKRYAVI 600
 :|||||:||:|||||:||:|||||:||:|||||:||:|||||:||:|||||:||:|||||:||:|||||:
 :||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:
 538 ysdtdgffatipgadetvkkkamefinyaklpaleleyegfkyrgfivtckkyavi 597
 Db 601 DEEGRITTRGLEVVRDWESEAKETOAKVLEALIKEDSVKEAVETVWDVSEIAKYQPL 660
 Qy 601 DEEGRITTRGLEVVRDWESEAKETOAKVLEALIKEDSVKEAVETVWDVSEIAKYQPL 660
 CC 598 deegkittgrleivrdwseaketgarvialealklgdvekavrikvekteklyevpp 657
 Db 661 EKLVHQITKDLSEKAIGPWAIAKRLAAGKIKVRGTTISYVLRGSKISDRVILL 720
 Qy 658 ekvlineqitrdkdykatgphavavkrlaargvirkptgtyisylkgsgrigdraipf 717
 Db 721 SEYDPKKHKKYDPYYENQVPAVLRLAEGYRKEDLYKQSSKQGLDAMIK 773
 :|||||:||:|||||:||:|||||:||:|||||:||:|||||:||:|||||:
 718 defdptkhkyaeyylenqvpaverlrlafgyrkdyqtrqvgisawlk 770
 RESULT 11
 ID AAW48453
 AC AAW48453;
 XX
 XX
 DT 15-JUL-1998 (first entry)
 DE Mutant KOD DNA polymerase; protein: 774 AA.
 XX
 XX
 XX
 OS Pyrococcus sp.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 311
 /note- "Changed from Tyr in wild-type
 to Phe in mutant"
 FT Domain 137..146
 /note- "Exo 1 region"
 FT Domain 206..222
 /note- "Exo 2 region"
 FT Domain 308..318
 /note- "Exo 3 region"
 PN EP822256-A2.
 PD - 04-FEB-1998.
 XX
 PR 24-JUL-1997; 97EP-0112760.
 XX
 PR 30-JUL-1995; 96JP-0200446.
 PR 29-JUL-1995; 96JP-0198911.
 PA (TOYM) TOYO BOSEKI KK.
 XX
 PT Imanaka T, Kamimura H, Kawakami B, Kawamura Y, Kitabayashi M;
 PT Komatsubarah, Takagi M;
 XX
 PS WPI; 1998-102909/10.

Modified thermostable DNA polymerase - with reduced 3'-5'
 PT exonuclease activity
 XX
 PS Claim 15; Page -; 54pp; English.

The invention claims a mutated thermostable KOD DNA polymerase of the present sequence. It has 0.01% of the 3'-5' exonuclease activity when compared with the wild-type KOD DNA polymerase. It also has a DNA extension rate of at least 30 bases/sec while retaining at least 60% of its activity after heating at 95 deg. C for 6 hours. The invention also

CC claims other mutated versions of the polymerase (AAW48449-WA8461). The invention shows that efficiency in DNA amplification can be achieved by using a mixture of two different DNA polymerases which are almost identical to each other with respect to thermostability and DNA extension rate but are different in their 3'-5' exonuclease activity. The mixture may contain the present polymerase or a mutant polymerase (AAW48454-WA8461), having 6-100% of the 3'-5' exonuclease activity of a native KOD polymerase. N.B. The present sequence is not given in the specification, but is made up using the given wild-type KOD polymerase sequence and the information provided in the claims section.

XX SQ Sequence 774 AA:
 Query Match 80.0%; Score 3221.5; DB 19; Length 774;
 Best Local Similarity 77.4%; Pred. No. 5.2e-243; Mismatches 598; Indels 3; Gaps 2;
 Db 1 miltdydtedgkpvirkfengefkkeydrtfepfyalkkddsaieevkitaerhg 60
 Qy 61 KIVRVDAVKKKFGLGRDVEWKLIFEHFEPDOPALRKREHPAVIDTYDIFAKR 120
 :|||:|||||:||:|||||:||:|||||:||:|||||:
 Db 61 tvvtvkrvekqkkgipewklyfthpdqvpairdkrhgavidiyedifakr 120
 Qy 121 LIDKGKIPMEEDEELKIMAFDIEFHEGDEFGKSIIMMSYADEEARYWTWNIDLPY 180
 ||||:|||||:||:|||||:||:|||||:
 Db 121 ldkglvpmedeeklmakrladietyhegefaegleilmlyadeegarrrtwkndlpv 180
 Qy 241 PKIHRMGSFAVEIKGRHIFDLPWVVRTINLPTILEAVNVEAVLGTKTKSAGBIAAT 300
 ||||:|||||:||:|||||:||:|||||:
 Db 239 pkiqrmqdrfavevkgrihfylpvirtrnlptyleaveavfqpaktyaetipa 298
 Qy 301 WETEESNMKLAQYSMEDARAYTELGEKFFPMEALAKLIGOSWVDSRSSSTGNLYEWYL 360
 ||||:|||||:||:|||||:
 Db 299 wetgenlervarfsmedadkvrlyelgekfplmeaqdlsrligslwdsrsstgnlvewfil 358
 Qy 361 RVAYERNELAPKDPDEBEYRRLRTYLGKVKEPERGLUNENITYLDFRCIYPSITVHN 420
 ||||:|||||:||:|||||:
 Db 359 ikayernelakpkdekelarr-rsgsyegryvkepergylmenivylfslsylsliith 417
 Qy 421 VSPDTLREGCKNVDAPTYKCFDGPSPITSPITGELITMROEKKKATIPIEK 480
 ||||:|||||:||:|||||:
 Db 418 vspdtlnregckeydavpqvghfrckdpfplslgdleerqkikknnatidpierk 477
 Qy 481 MLDYRQRAVKLUHANSYGMGPKARWYKSCAESVTAWGRHYIEMTIKEEKGCFKL 540
 ||||:|||||:
 Db 478 lldyraqraikllansyggryarwyckeasvytawgryeitntikeekygfki 537
 Qy 541 YADTGGYATPGKEPKETIKKKAKERFLKYNSKPLGPLELEYEGFLRGFVAKRYAVI 600
 :|||||:||:|||||:
 :||:||:||:||:||:||:||:||:||:||:||:||:||:
 538 ysdtdgffatipgadetvkkkamefinyaklpaleleyegfkyrgfivtckkyavi 597
 Db 601 DEEGRITTRGLEVVRDWESEAKETOAKVLEALIKEDSVKEAVETVWDVSEIAKYQPL 660
 Qy 601 DEEGRITTRGLEVVRDWESEAKETOAKVLEALIKEDSVKEAVETVWDVSEIAKYQPL 660
 ||||:|||||:||:|||||:
 Db 598 deegkittgrleivrdwseaketgarvialealklgdvekavrikvekteklyevpp 657
 Qy 661 EKLVHQITKDLSEKAIGPWAIAKRLAAGKIKVRGTTISYVLRGSKISDRVILL 720
 Qy 721 SEYDPKKHKKYDPYYENQVPAVLRLAEGYRKEDLYKQSSKQGLDAMIK 773
 ||||:|||||:
 Db 718 defdptkhkyaeyylenqvpaverlrlafgyrkdyqtrqvgisawlk 770

RESULT 12

XX	Sequence	774 AA;	Score	3218.5;	DB	19;	Length	774;	
Best Local Similarity	77.4%	Pred. No.	9e-243;	Mismatches	90;	Indels	3;	Gaps	2;
Matches	598;	Conservative	82;						
Qy	1	MIFDDTYITKDGKPIRIFKENGKPIEKLDPHFOQPYIALLKDDSAIDEIKATGERHG	60						
Db	1	mildtayitegdgkpvrifkengerkleydtfepyfvalkksaeevkkitaerhg	60						
Qy	61	RIVRYVDAVKRKKEGRDDEVWKLIFEFHPODVPALRKTRHPRAVIDYEDPFKRY	120						
Db	61	tvvtvkrvekkqkfkgrpewvkwifthpdvpairkirkrengavidiyedpfkry	120						
Qy	121	LIDKGKIPMGSDEELKMAFDIETFYHEGBEGKREBIMSYADEEARVITWNIDPY	180						
Db	121	11dkglvpmegdeelkmlafnetlyhegefaeagpilimisyadegarvitwkhdpv	180						
Qy	181	VDVVSNEREMKRFQIVREKDPDVILTYNGDNFOLPYLIKRAEKIGVILLGRDKEHPE	240						
Db	181	vdvvssteremikrlfvwvkekdppdityngdnfotayikrccklginalgrgs-e	238						
Qy	241	PKIIRNGDSFAVEIKGRIHEDLPVWRRTNLPTTLEAVYEAVGRTSKLGKBEIAAT	300						
Db	239	pk1qmgdrfavevngrihelyp1lrltlnlpttleaveyavfqpkckvyeitpa	298						
Qy	301	WETEBSMKKLAQYSMDARATYELCKEFFPMEALAKLIGOSWMSVRSSTGNLVEWYL	360						
Db	299	wetgenlervarvnsnedakvtyelkefipmeaqslrlqslwvrsstgnlvefil	358						
Qy	361	RVAYEENELAPNPKPDEEEYRRRLRTYLGVVKKEPERGLMENTVLDERLYPSTIVTHN	420						
Db	359	rkayernelapnkpdekelarr-rasvyeggyvkepergwienvyldtslypsiilthn	417						
Qy	421	VSPDTLEREGSKNQYDVARINGYKFKDFPERIPSILGEUTMRQIKKKMATTPIEK	480						
Db	418	vspdtlinregckeydvarpgqyghrfckdfpgfipslgdlleertgkikkatdpielk	477						
Qy	481	MLDYRORAKVILHANSVYGMYPKARWYSKCAESVTAWGRHYITMVKIEEFGFKVL	540						
Db	478	ldlyqraikilanasyyygqyarrtwycrcasvtaqreyimtikieekyfki	537						
Qy	541	YADTCGFTATPGKEPTEIKKAKRPLKVNNSKUGGLELEYEGSYLRFGRFVAKRYYAVI	600						
Db	538	ysdtogffatipgadetaetvkkkamafinlyinaklgaleleygykrgfvtkkyyavi	597						
Qy	601	DEEGRTTRGELEVWDSEIAKEQAVKIALEALKEDSVKAVEWKDVVEELKQVPL	660						
Db	598	deegkittrggleivrvrdwseiakeqavkialealkgdvkaevkivkevtekskyevpp	657						
Qy	661	EKLVHHEQITKDSEYKAICPHVATAKRLAAGKIVKRPGTISYVLRGSGKIDRVL	720						
Db	658	eklvineqitkdklykatphavakriaargvirkpvisyvikksgridraip	717						
Qy	721	SEYDKRKHKYKDPDYIENOLPAVWRILEFGYRKEKDLYKFOSSKGYGLAWLK	773						
Db	718	defdptkhkyaeyylenqvlpavervlrafgykedlryqktrqvglsawlk	770						

Search completed: May 28, 2002, 14:18:46
 Job time: 104 sec

Q9HH98	PRELIMINARY;	PRT;	759 AA.
Q9HH98;			
AC			
01-MAR-2001 (TREMBLrel. 16, Created)			
01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
01-MAR-2001 (TREMBLrel. 19, Last annotation update)			
DE	DNA POLYMERASE (FRAGMENT).		
GN	POL.		
OS	Pyrococcus sp. (strain ST/00).		
OC	Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.		
OX	NCBI_TAXID=69015;		
RN	SEQUENCE FROM N.A.		
RC	STRAIN=ST/00;		
RA	Querellou J.-J.E.; Cambon M.A.; Lesongeur F.; Barbier G.;		
RT	"Thermococcales taxonomy and phylogeny based on the comparative use of 16S rRNA, 16S-23S rDNA intergenic spacer and family B DNA polymerase genes";		
RR	Submitted (OCT-1999) to the EMBL/GenBank/NCBI databases.		
RL	- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE + DNA(N).		
CC	-1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.		
DR	EMBL; AJ250332; CAC12847.1; -.		
DR	HSSP; P56689; ITGO.		
DR	InterPro; IPR02064; DNA_POL_B.		
DR	InterPro; IPR01648; Ribosomal_S18.		
DR	Pfam; PF00136; DNA_POL_B; 1.		
DR	Pfam; PF03104; DNA_POL_B_exo; 1.		
DR	PRINTS; PRO0106; DNAPOLB.		
DR	SMART; SM00486; POLB; 1.		
KW	PROSITE; PS00116; DNA_POLYMERASE_B; UNKNOWN_1.		
FT	DNA replication; DNA-binding; DNA-directed DNA polymerase.		
NON-TER	759 AA; 88431 MW;		
SEQUENCE	A8459B6AAD8F3AF3 CRC64;		
CC	Best Local Similarity 74.7%; Score 3081; DB 1; Length 759; Matches 569; Conservative 99; Mismatches 90; Indels 4; Gaps 3;		
QY	1 MTFDTDVITDKGKPILIRIFKKNGEKKFIELDDHFQPYIYALLKKDSADEIKAIGERHG 60		
Db	1 MILDADYITENKPKIIRFKKKENGKFVYEDNFRPYIYALIKDSDAIDDVKITSERHG 60		
OY	61 KIVRUVDAVKVKKFLGRDVEWKLIFEHPODPALRGKTRHPAVIDEYDIPFAKRY 120		
Db	61 KVVRVIDEVKVSKKFLGRPIEVWKLIFEHPODPALRGKTRHPAVIDEYDIPFAKRY 120		
QY	121 LIDKGKIPMGEDEELKLMADFETFYHEGDEGKGETIMISTADEEVARVITWNKIDPY 180		
Db	121 LIDKGKIPMGEDEELKLMADFETFYHEGDEGKGETIMISTADEEVARVITWNKIDPY 180		
OY	181 VDVVSNTREMIKRFVQITVREKDDVITYNGNDFLPYLIKRAEKGTVLGLRDKHPE 240		
Db	181 VDVVSNTREMIKRFVQITVREKDDVITYNGNDFLPYLIKRAEKGTVLGLRDKHPE 240		
OY	241 PRTHRMGDSFAVEIKGRHEDLFPWVRTINLPTYTLAEAVYAVLGKTSKLGAEETAI 300		
Db	239 PRMQRIGESLAVEIKGRHEDLFPWVRTINLPTYTLAEAVYAVLGKTSKLGAEETAI 298		
OY	301 WTEESMKKLAQYSMEDARAYELGKFFPMABEALAKLIGQSYWDVSRSTGNLVEWLL 360		
Db	299 WETGKGULPMEGNEEFLSFLAVDITLYHGEEGEKGPKPIMIYADEGAKVITWKDILPY 358		
OY	361 RYAYERNELALPKPDKDEEYRRRLTGYYKEPERGLWENNTYLDPRCLPYSITVN 420		
Db	359 RYAYERNELALPKPDKDEEYRKLRRESTEGGYKEPEKGWLWEGIVSLSRFLSPSITHN 418		
OY	421 VSPDTLREGCKKNYDVAPIGYFKCKDFGFPFISLGEGLITMROEIKKKMKTADPTEK 480		
Db	419 VSPDTLNREGCKGYDEAPEGVHFRCKDFGFPFISLGEGLITMROEIKKKMKTADPTEK 478		
OY	481 MARYDORAVKLHANSYYGGMGPKARRYKCEGAESVTAAGRAHYIEMTKEKEFGKVL 540		
RC	- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE + DNA(N).		
CC	-1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.		
DR	EMBL; AJ250332; CAC12847.1; -.		
DR	HSSP; P56689; ITGO.		
DR	InterPro; IPR02064; DNA_POL_B.		
DR	InterPro; IPR01648; Ribosomal_S18.		
DR	Pfam; PF00136; DNA_POL_B; 1.		
DR	Pfam; PF03104; DNA_POL_B_exo; 2.		
DR	PRINTS; PRO0106; DNAPOLB.		
DR	SMART; SM00486; POLB; 1.		
KW	PROSITE; PS00116; DNA_POLYMERASE_B; 1.		
FT	DNA replication; DNA-binding; DNA-directed DNA polymerase.		
NON-TER	781 AA; 90384 MW;		
SEQUENCE	9B2570EAB30C372C CRC64;		
CC	Best Local Similarity 36.6%; Score 1204.5; DB 1; Length 781; Matches 294; Conservative 141; Mismatches 282; Indels 87; Gaps 21;		
QY	2 IFDITYTIDKGKPTRIFK-KENEFKFFILDPHCPYIYALLKKDSADEIKAIGERHG 58		
Db	7 ILDSSYDVENPKPYIYIWIDKEGRIVVLUKEKKFPPYFVALVDDSYNIDIRKETILSK 66		
OY	59 HKIVRUVDAVKVKKFLGRDVEWKLIFEHPODPALRGKTRHPAVIDEYDIPFAK 118		
Db	67 PYSPTISIDY--EERKGYFSPVVKIETVPAVVRVYDEKAVIKGVSYLEADIRFV 124		
OY	119 RYLIDKGKIPM-----EGEB-----ELKLMADFETFYH 147		
Db	125 RYSIDINLKPFWTAEVEIKENFRVKKVYELKINKLYEDKPEKLVAFDIEVY-- 182		
OY	148 EGDFG-----KGEIMISYADEEVARVITWNKIDPYVWVWSNREMIMKREYQIVREKD 202		
Db	183 --NQGSPMRRDPDIIIGWATKEGKQFLADKVN-----DLRAFTEEFNIVQTYD 231		

QY	[1] SEQUENCE FROM N.A.
QY	625 RYVGFLIEDGRIDIVGFFAVRGWCWELAKEVQEKAAEIVLNTGNDKALSYIYKVIKOLRE 684
Db	STRAIN=DSM2705;
QY	MEDLINE=9523290; PubMed=7721707;
RA	Uemori T., Ishino Y., Doi H., Kato I.;
RT	"The hyperthermophilic archaeon Pyrodictium occultum has two alpha-like DNA polymerases.";
RT	J. BACTERIOL. 177:2164-2177(1995).
CC	- - CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N
CC	YROPHOSPHATE + DNA(N).;
- - SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.	
DR	EMBL: D8874; BAA0580.1; -.
DR	HSSP: P56689; 1TGO.
DR	InterPro: IPR02064; DNA_Pol_B.
PFam	PF00136; DNA_Pol_B; 1.
DR	PF03104; DNA_Pol_B_exo; 1.
DR	PRINTS; PR00106; DNAPOLB.
DR	SMART; SM00486; POLBC; 1.
KW	PROSITE; PS00116; DNA_POLYMERASE_B; 1.
KW	DNA replication; DNA-binding; DNA-directed DNA polymerase; Nucleotidyltransferase; Transferase.
SEQUENCE	803 AA; 92656 MW; E9C092F26A8D23FA CRC64;
QY	Query Match 29.6%; Score 1190; DB 1; Length 803; Best Local Similarity 33.7%; Pred. No. 8.2e-62; Matches 281; Conservative 164; Mismatches 278; Indels 110; Gaps 23; Gaps 23;
QY	1 MIFDIDY-----ITKGKPIIRFKKENGEFKIEELDPHQFQPYTIALKD-D 45
Db	8 VLLDSSYEELGKERVWLMGTLDGCKRVL-----LDHFRFRYFYIALRGVE 55
QY	46 SAIDEIKAKGKGERHKIVR--WVDAVKVVKELGRDVEWKLIEHPOVPAKGKIREH 103
Db	56 DMVEETAA--SIRRLSVWVKSPIDAKPLDKRFRGPRKRKAVITMIPESVRHYRAVKI 113
QY	104 PAVIDIYEDIPFAKRYLIDKL-----IPME-----GDEE----- 134
Db	114 EGVEDSLEADIRFAMYRIDLKRLYPFTVVRIPDAGRNPGRFVDRVYKAGDKEHEPEPKIHR 173
QY	135 -----LKLMAFDIETFYHEGD-EFGKEIIMSYADEE-EARVITWNIDLPVYDVVS 185
Db	174 TRIDLPPMPLVVAFFIEVNRGSGSNPARDPVIIVLSDRGEKERLIEAEGHD----- 225
QY	186 NEREMIKREFOVYIREKDPPLVITYNGDNFDLPLVLIKRAKLGVYLLGKDEHEPEPKIHR 245
Db	226 -DRRVLRFLFVEYVRAFDPDLIVGYNNSHFDWPLYMERAKRLGKIDVTR-RVGAEPPTSV 283
QY	246 MGDSFAVEKGRHIFDFFWVRRVNLNPFTYLEAVEAVALGKTKSKLGAEE--IAINE 302
Db	284 YGH---VSVQGRNLVYDYLAYEEMPEIKKTLTEVAEYLGVMKSERVIEWMWRPEYND 340
QY	303 TEESMKKLAQYSMEDARATYELGEKEFPMAELAKLGSVWDYSSRSSTGNLVEWLLRV 362
QY	311 DEKKRQLLERVALDVRATYGLAEKMLPFPALOSVTGFLDOVAGMGFRLEWLMRA 400
Db	363 AYERNELAPNKPDPSEEYRRLRTYLGGYKEPERGLWENITYLDFRCIYPSITVHNS 422
QY	401 AYDMNELLVNPNRV-----RGRGESYKGVAVLKLPLKGVHENVVWVDFSSIMPSIMKYNG 454.
QY	423 PDTL--ERRGCKNID--WAPIVGKFKCOPGPGIPIPSIIGELITMRQEKKKMA-TIDP 476
Db	455 PDTVDPSPCPKPGGYVAPEVGHFRASPPGPFKTVLNLKRRQVKEMKEFPPS 514
QY	477 TEKKMDYQORAVKLHANSVYGYMCKARWYKSKCAECASTWAGHYIEMTIKEEFG 536
Db	515 PEYRIDEQSKQKALVLYANSVYGYMCKARWYKCAEATWAGRNLTIAIEVARKLG 573
QY	537 FKVLYADTGFYATIPGEKEPTEKKAKELFYINSKPLOLELEYEGFYLGRGFFV-ACK 595
Db	574 LKVIYGDTSFLVWYD-----KEKVKELIEFVEKEL--FEIKDITYKKVFTEAKK 624
QY	595 RIAYIDEDEEGRITRGELEVVRDWSLIAKETQAKLIAKEDSVEKAVELVKDVVEELAK 655
QY	RESULT 7
Q9P9M2	PRELIMINARY; PRM; 785 AA.
RP	SEQUENCE FROM N.A.
RC	STRAIN=DSM2705;
RX	MEDLINE=20100754; PubMed=10633098;
RA	Kaehler M.; Antranikian G.;
RT	'Cloning and Characterization of a Family B DNA Polymerase from the Hyperthermophilic Crenarchaeon Pyrobaculum islandicum.';
RT	J. BACTERIOL. 182:655-663(2000).
CC	- - CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N
CC	YROPHOSPHATE + DNA(N).
- - SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.	
DR	EMBL; AFI95019; RAF27015.1; -.
DR	HSSP; P56689; 1TGO.
DR	InterPro: IPR02064; DNA_Pol_B.
DR	PFam; PF00136; DNA_Pol_B; 1.
DR	PRINTS; PR03104; DNA_Pol_B_exo; 2.
DR	SMART; SM00486; POLBC; 1.
DR	PROSITE; PS00116; DNA_POLYMERASE_B; 1.
KW	DNA replication; DNA-binding; DNA-directed DNA polymerase.
SEQUENCE	785 AA; 89748 MW; 10FB8B6A8F3730D CRC64;
QY	Query Match 27.9%; Score 1125; DB 1; Length 785; Best Local Similarity 33.8%; Pred. No. 5.3e-58; Matches 273; Conservative 156; Mismatches 284; Indels 94; Gaps 23; Gaps 23;
QY	4 DMDYITKGKPIIRFK-KENGBEKFIEELDPHQFQPYTALKDODSAIDBEIKAKGKGERHKI 62
Db	9 DITIYAVGSVPEIRFGLLSSGERVVLIDRSKPKFYV--DCAVCEPAALK-TALSRV 63
QY	63 VRVDAVKVKKFLGROVEWKLIEHPOVPAKGKIREHPEPKIHRMGDSFAVEIKGRHIFDPLPVVRR 122
Db	64 APIDDVQVERPFLGRSKKFLKVIAKIPEDVVRKLRREAMSTIPRVSGVYEAIDIRFYMRYI 123
QY	123 DKGLLIPM-----EG-----DE---ELKLMADIETFYHEGD-EF 152
Db	124 DMGVFCPSWNVAAVEEGGRGCGIPPTVVSQWVGDSEPPSLKVMADFDEVINERGSPD 183
QY	153 GKETIMSYA---DEEAVRITWNIDLPVYDVVSNEREMIKRQVIREKODPVVLY 209
Db	184 IRDPVWMAIKTNDGHFVFEASKG-----DGRGVRAFDVIRTSDDPDPVGY 232
QY	210 NGDNFDLPLVLIKRAEKGIFTGLLGRDKHEHPEPKIHRMGDSFAVEIKGRHIFDPLPVVRR 269
Db	233 NSNGFDWPLVYERAKAVGVPLKVDRSLNSPQOOSVYGHW---SIVGRANVNDLYNTVEEF 287
QY	270 INLPFTYLEAVEAVALGKTKSKLGAEE--IAINE 325

Db	891 GOTQTGL 897	Db	348 KI---PLDDVANVGTSNVDSLIRADRENIVGPMQH-----IKTERIQQGVHST 398
RESULT	11	QY	396 ERGLWENITYLDFRCLPSPSIYVTHNPSDTEREGCKNYDAPIVGKFC--KDPGFP 453
09HJRO		ID	09HJRO PRELIMINARY; PRT; 796 AA.
		AC	09HJRO: 01-MAR-2001 (TREMBrel. 16, Created) 01-MAR-2001 (TREMBrel. 16, Last sequence update)
		DT	01-DEC-2001 (TREMBrel. 19, Last annotation update)
		DE	DNA POLYMERASE (POLB), LARGE CHAIN RELATED PROTEIN.
		GN	TAA097.
		OS	Thermoplasma acidophilum.
		OC	Archaea; Euryarchaeota; Thermoplasmales; Thermoplasmataceae;
		OC	Thermoplasma.
		OX	NCBI_TAXID:2303;
		RN	[1]
		RP	SEQUENCE FROM N. A.
		RC	STRAIN=DSM 1728;
		RX	STRAIN=DSM 1728; PubMed=11029001;
		RA	Ruepp A., Graml W., Santos-Martinez M.-L., Korette K.K., Volker C., Mewes H.-W., Frishman D., Stocker S., Lugas A.N., Baumeister W.;
		RT	"The genome sequence of the thermoacidophilic scavenger Thermoplasma acidophilum."
		RL	Nature 407:508-513 (2000).
		CC	-I- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE + DNA(N).
		CC	-I- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
		DR	EMBL: AL440655; CACI036.1; -.
		DR	InterPro; IPR02064; DNA_Pol_B.
		DR	InterPro; IPR00077; FGGY_kin.
		DR	pFam; PF00136; DNA_Pol_B; 1.
		DR	PRINTS; PRO0106; DNAPOLB.
		DR	SMART; SM00486; POLBC; 1.
		DR	PROSITE; PS00933; DNA_Polymerase_B; 1.
		DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
		KW	Complete proteome; DNA replication; DNA-binding; DNA-directed DNA polymerase.
		SQ	SEQUENCE 796 AA; 91210 MW; DBCD80F143EB929 CRC64;
		QY	Query Match 15.5%; Score 623; DB 17; Length 796; Best Local Similarity 25.7%; Pred. No. 1.9e-28; Matches 211; Conservative 161; Mismatches 314; Indels 134; Gaps 31;
		QY	16 IRIF-KKENGKKEIPLPHFQVYIYALKDSAAIDEKAIRGERHKGKIVRUDAVAVKK 74
		QY	19 VELFGRREGESVAALYFGFRPY-YDVEPBEA--LKVIOND-----PEFYKMEDK 67
		QY	75 FL---GRDVEWKKLIFHPDQVALRKIREHAPVIDEYDIFPKAKYRILDKG----- 126
		QY	68 RLWIRGKENVKRYIYRSPWKVPEYRK----CPFEVLAADIPIFHRIYDLDGSCVK 122
		QY	127 -----IPMEDI-----ELKIMAFD-ETFYHEGDEGKGKELIMIS 161
		Db	123 IIGENTSDRETSFTTDIVADR1ENYNDENFNPKVLSFDVENEINRENVEDYKGKLVIG 182
		Db	162 WADBEARVITWKNIPLPVIVSNEREMIRPVQVREKPDPLVITYGNDPLYK 221
		QY	183 YSVSFQCKTVIG-----SLSGEQDILSFDLRAEPDPLVITYGNDYDVK 234
		QY	222 RAELKGVYLLGRDKHEPEPKTHRMDSFAVEIKGRIHDFLPVWRRTINLPTYLEAVY 281
		Db	235 RMDYRIGKLEIIGRDSIP---RIMQF-WRVRGLISDWTWMSVRLHPKESLYVA 289
		QY	282 EAVLGKVKSKGAAEELNAAWTEESMKLQAOYMSMEDARATELGKEFPME---AELA 336
		Db	290 NMLLGECKDNIDRLHEDW--KRRREVIAYCICKDADLTTRIFEEKLMVMRLMMSSVT 347
		QY	337 KUIGQSVWDVRSRSTGNLVEWILLRVAYERVELAPNPKDEEYRRRLRTYL-GGYVKEP 395
		QY	Query Match 15.5%; Score 622.5; DB 17; Length 800; Best Local Similarity 25.9%; Pred. No. 1.9e-28; Matches 215; Conservative 165; Mismatches 306; Indels 145; Gaps 33;

QY	9	TRDGKPTIRFKKENGKFTIELDPHR-----QPYIALLKDSAT--DEIKATKGERH	59	RX	PubMed=11572479;
Db	27	SREGKSVSALF-----FGKPKFDVVERPNEVLLSIRNNDDEVKEEDKVWLWVGSMH	78	RA	Kawarabayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
Qy	60	GKIVRVVDAVKKKELGRDVEUWKLIFERPODVAALRGKTRERPAVIDYDIPFAK	119	RA	Seike M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
Db	79	-----NVKRIYIRSPWKVPEYR--RRCP--FEVLADEPFHRR	112	RA	Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamaya M., Kato Y.,
Qy	120	YLIDKGL--IPMEGDE-----	113	RA	Yoshizawa T., Tanaka T., Kudo Y., Yamazaki J., Kushida N., Oguchi A.,
Db	113	FYDPEGLACVTRGIDISCTEGNETDLYLKDIENTPDFINVNUKVISFDVENEINRE	172	RA	Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamashita A.,
Qy	210	NGDNFDLPLYLKRAEKGIVLGLRDKHEPEPKLIRMGDSFAVEKGRILDFLFVWR	209	RA	Oshima T., Kikuchi H.,
Db	173	NVEDIGKILWIGS-----IMFQDQ-KKGEGHDEKEILYRFDLIRADPQVITY	224	RA	"Complete genome sequence of an aerobic thermoacidophilic
Qy	225	NIDGYDMLVQIQRMEQYGIHLNLGRGSVP---RRIMDF-WRVHGRLLSDTAWNWKI	279	RT	Crenarchaeon, Sulfolobus tokodaii strain7.";
Db	270	INLPTVLEAVYEAVLGKTKSKLGAEELAIWIETEESMKLQYSMEDARATYEGKEF	329	RL	DNA, Res. 8/12/3-14/0/2001';
Db	280	LHPKHSLSLDAKMLLGEKDSDRNTLEEWKQD--EVITSYCIKDADLTFLRFEKL	337	DR	EMBL: AP000986; BAB6493.1; "-DNA-directed DNA polymerase; Hypothetical protein; Complete proteome
Qy	330	PME----AEIYLKLGQSVMWRSRSTGQNLVWYLLRVAVERNELAPN---KPDEEY	379	SQ	SEQUENCE 876 AA; 101584 MW; C013AA0FB0831F96 CRC64;
Db	338	VLERMEMSTWTKL--PLDVAAGTSNTSYVDSLIRADRENINGVPMNAHNLKDEIQ	393	Query Match	15.3%; Score 616.5; DB 17; Length 876;
Qy	380	RRRLRTYLGCVYKEPERGLWENTYLDFLRCLYPSITVTHNVSPTLEREGCKNVDAP	439	Best Local Similarity	26.6%; Pred. No. 4.8e-28;
Db	394	394 -----GGVHHTIGAGLYSNVLDFKSMYPSMIIKVNCFITLDPKG---EVSP	440	Matches	231; Conservative 140; Mismatches 316; Index 181; Gaps 35;
Qy	440	VGYKFC--KDPGFPSPISLGLPLTHQEIKKKMKTIDPEKMDYQRRAVKHANSVY	497	Db	53 LLQVYDGGKKSKAVCKLKDKEKKYLYLQYDNGHKPF-----LTDIDPERV 100
Db	441	TGIRETSPEKKKGLPILOBLLMADDDVVKRMLAKSEDERLYFDGQNAKVNLT	500	Qy	61 KIVRYV-----DAVKYKKFGLRDVWKLIFEHFHDQDVPLRGKTRHPEAVIDE 111
Qy	498	GYMGYPKARMYSKECAESVTAWGRHYIEMTKEEKGKVLKADTGFYATIPGEKPE	557	Db	101 KIPKVVRDPSFDHLETVIKI-DPSGNKIKLTQKIVVKDPLAVRMRNSPK-----AYE
Db	501	GYLASSFYRFDPDKTSATTAFARETIKHI-DVLESGHRYVQDTSFVESDVRQ	559	Qy	112 YDIPPAKRYLIDKGLP-----MEGDE-----
Qy	558	TIKKAKEFLKYINSKLPLG--LELEYEFPYLRGFFVAKKRY-----VIDEEGRITR	609	Db	154 AHKTFNNYIYDGLPQYVWVKGKLGQYRPELKEDEIRKAFAASDEMTEKAVND
Db	560	DAIKIGKELSKL-SEEGLTLDQVNLDPFFSHG--AKKRYAKSCVKVIFEDMKGEIK	615	Qy	134 -----EKLMAFDIETFY-----HEGDFGKELIMSYA-DEEAVRIVWK
Qy	610	GLEYERWDWSLAKETOAKVLEAIKEDSYEKAELVKDVYET--AKYQVPLEKVIHE	667	Db	214 WIPIFSEVSDPKVRAIDEVYTPIKGRIPDPEKEF--PISISLAGNDGTRVYL
Db	616	GYEVRENDSEFQDQSTALKSVDLFDLDRD-VQGAVNADYVADKVKVNRGDPDSITELSVIR	674	Qy	112 YDIPPAKRYLIDKGLP-----MEGDE-----
Qy	668	QITKDSEYKAIGPHVA--IAKRLAAKGKVRPGTISIVLRSGGKTSDRVILSEYD	724	Db	175 NIDL-----PYDVWSNEREMKRFVQIREDPDYLITYGDNFDLPLYKRAEK
Db	675	675 TV-KDPSYKANTDSLANTRAARKLJLGERGTFPQEMKVSWIT--NGKKUPQEVETYG	731	Qy	271 REDVNSQITKHDVIVETFSKSERELIRRFDTI--LDYPTILTFNGDDFLPYTYRALKL
Qy	725	PK-KHXYDPDYYIENQVLPAVLRLIAFGYRKEDLYQSSKQVGLDAWKK	774	Db	227 GVTLLIGRDEHEPEPKIHRMGDSFAVEKGRIFDLP-WVRTI-----NLPYTYE
Db	732	732 SKLESKPDWDXAK RLSETINRVIDVF--RKD--MVQGNRRISSADSFSK	777	Qy	329 NFT-----PEPFDIENDSKYLA-----GHDLYKPFENRATRNAYFEGKYNED
RESULT	13	Q971C7	PRELIMINARY; PRT: 876 AA.	Db	379 AVYEAVLGLKTKSKLGAEELAIWIETEESMKLQYSMEDARATYEL--GKEFPME
ID	0971C7	AC		Qy	400 VGYKCKDFCFFPSPISLGLPLTHQEIKKKMKA-TIDPEKMDYQRRAVKHANSVY
DT	01-DEC-2001 (TREMBLrel. 19, Created)	Qy	498 GYMGYPKARMYSKECAESVTAWGRHYIEMTKEEKGKVLKADTGFYATIPGEKPE		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)	Db	609 GVFGAENFPLYPAYAESAESVTAWGRHYIEMTKEEKGKVLKADTGFYATIPGEKPE		
DE	PUTATIVE DNA-DIRECTED DNA POLYMERASE.	Qy	558 TIKKAKEFLKYINSKLPLGLEYEFGYRGEFVA--KRYAVIDEGRITRGLV		
GN	ST1426.	Db	664 -KEKLEITKFKVKGFG--LDLEVDYK--FVAFSLKKNYQVPGKTDQKGM		
OS	Sulfolobus tokodaii.	Qy	614 VRDMSEIKAEGVHATKLAAGKIVKRPGTISIVLRSGGKTSDRVILSEYD		
OC	Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.	Db	717 KKRNPTEFKEFNEVQKLVFTTNSPDDPKIRDQLEYKTEYKLRGYNLDELA		
OX	NCBI Taxid=11955;	Qy	667 EQITKDSEYKAIGVHATKLAAGKIVKRPGTISIVLRSGGKTSDRVILSEYD		
RN	[1]	SEQUENCE FROM N.A.	RC	STRAIN-JCM 10545 / 7;	

QY	569	YINSKPLGLLELEFGYFLGGFFVAKKRAV-----DEBRITRGLEVRROSEIA	622
Db	781	WWSGFSPSPRLEREEFKVYFFYLJLSSRDPADHRMOCKGLEAVRROCNPLV	840
QY	623	KETQAKVLEALKERVKERAVETIVDVEETAKYQPLEKVLVHQITKULSEYKAGPH	682
Db	841	ANLUTASLRRLIDRPEGAVAHQDVISDLCNKRIDISOLVITKELTRASDYACKQAH	900
QY	683	WAIKRLAANGIKVPR--GTTISYVTLRSG-----KISDRVILSEKUPKKHHKDVPY	734
Db	901	VELAERMRKRDGPSAFLGLDRVYVIIASAARGVAAYMKSDFLVLEHSI,P---IDTOY	956
Qy	735	YIENQVLPVALRILE	749
Db	957	YLEOOLAKPLRLIFE	971

search completed: May 28, 2002, 14:21:54
job time: 257 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5						
copyright (c) 1993 - 2000 Compugen Ltd.						
M protein - protein search, using sw model						
run on: May 28, 2002, 14:18:07 ; Search time 14.97 Seconds						
scoring table: BLOSUM62 (without alignments)						
gapopen 10.0 , Gapext: 0.5						
total number of hits satisfying chosen parameters: 105224						
minimum DB seq length: 0						
maximum DB seq length: 200000000						
post-processing: Minimum Match 0%						
database : SwissProt_40:*						
predicted No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.						
SUMMARIES						
result No.	Score	Query	Match	Length	DB	ID
1	346.5	85.6	1829	1	DPOL_THEST	033845
2	3356	83.4	1702	1	DPOL_THELLI	P30317
3	3235.5	80.8	773	1	DPOL_THEGO	P56589
4	3204.5	79.6	775	1	DPOL_THES9	Q56366
5	3164.5	78.6	771	1	DPOL_PYRAB	P77916
6	3163.5	78.6	771	1	DPOL_PYRSE	P77932
7	3151	78.3	775	1	DPOL_PYRFU	P80661
8	2934.5	72.9	1312	1	DPOL_PYRSF	Q51334
9	2868	71.7	1235	1	DPOL_PYRHO	059610
10	2809	69.8	1523	1	DPOL_THEME	P74918
11	2782	69.1	1671	1	DPOL_PYRKO	P77933
12	2769.5	68.8	1699	1	DPOL_THEGB	Q91nb84
13	2535.5	63.0	1668	1	DPOL_THEHY	Q91nb05
14	1392	64.6	824	1	DPOL_METVO	P52025
15	1301	32.3	781	1	DPOL_ARC25	Q29753
16	1292	32.1	1634	1	DPOL_METJA	Q58295
17	1299.5	58.6	1	DPOL_METH	Q27276	
18	1138	28.3	784	1	DPO2_AERPE	Q93746
19	697.5	17.3	1105	1	DPO2_ORYSA	Q91re6
20	693.5	17.2	1088	1	DPO2_SOYBN	Q48901
21	691	17.2	763	1	DPO3_SULSH	P05706
22	676.5	16.8	764	1	DPO3_SULSO	P95979
23	675.5	16.8	1081	1	DPOD_ARATH	Q91vn7
24	657.5	16.3	1038	1	DPOD_CANAL	P46588
25	657.5	16.3	1086	1	DPOD_SCOPA	P30316
26	624.5	15.5	1092	1	DPOD_DROME	P54358
27	621.5	15.4	1097	1	DPOD YEAST	P14436
28	619.5	1103	1	DPOD RAT	Q54747	
29	616.5	1081	1	DPOD_SULPH	Q50607	
30	616.5	15.3	882	1	DPOD_SULSO	P26611
31	616.5	1105	1	DPOD MOUSE	P54341	
32	615.5	1103	1	DPOD_MESU	P92831	
33	614.5	1094	1	DPOD_PLAFK	P30315	

FT	CHAIN	1393	1441	DNA POLYMERASE, 3RD PART (POTENTIAL).
FT	CHAIN	1442	1598	INTEIN III.
FT	CHAIN	1599	1829	DNA POLYMERASE, 4TH PART (POTENTIAL).
SQ	SEQUENCE	1829 AA;	211875 MW;	A113A8BC57EB9CB3 CRC64;
Query	Match	85.6%	Score 3446.5;	DB 1; Length 1829;
Best Local Similarity	42.2%	Pred. No. 3.2e-186;	Matches 771;	Conservative 0; Mismatches 3; Indels 1055; Gaps 3;
Qy	1	MIFDTDYITKDOCKPIRIFKKNGEKFKEILDPHFQPYIYALIKDSADEKAIKGRRHG	60	Qy 494 ----- 493
Db	1	MILDTDYITKDOCKPIRIFKKNGEKFKEILDPHFQPYIYALIKDSADEKAIKGRRHG	60	Db 1021 GMLRTLKWIFGEESKRIRTFRNRYLFHLEELGFVKKLPRGYEVTDWEGLKRYRQLYKLK 1080
Qy	61	KIVRVVDAVKVKFKLFLGRDVEWVKLFLFHEDPQDPAVLRGKIBHBAVIDYEVDPFACKY	120	Qy 494 ----- 493
Db	61	KIVRVVDAVKVKFKLFLGRDVEWVKLFLFHEDPQDPAVLRGKIBHBAVIDYEVDPFACKY	120	Db 961 KEIWELEVVKVGLLIVPKVKVKEAVINPELISKLPDADTAVWMTPVKGRKNEFK 1020
Qy	121	LIDKGKLIIPMGEDEELKMAFDOTETHEGDERGKGETIMESYADEESEARVITWKNDLPY	180	Qy 494 ----- 493
Db	121	LIDKGKLIIPMGEDEELKMAFDOTETHEGDERGKGETIMESYADEESEARVITWKNDLPY	180	Db 1081 NERYNGNKRKEVLVRENDIKDSVSCFPRKELERWKTGTXKGFRRXKCILKVIDDFGKFLGYY 1140
Qy	181	VDVVSNEREMIKRFVQVREKDPDVITYNGDNFDPLYLIKRAEKGVTLIGDRKEHPE	240	Qy 494 ----- 493
Db	181	VDVVSNEREMIKRFVQVREKDPDVITYNGDNFDPLYLIKRAEKGVTLIGDRKEHPE	240	Db 1141 VSEGYAGAQKNKTGGMSYSVKLYNENPNVLDKMDKNTAEKFFGKVRVGKNCVDPKRMAYL 1200
Qy	241	PKHMGDSFAVEIKGRIHDFPQVVRITNIPFTYLEAVEAVLGKTSKIGAEETAI	300	Qy 494 ----- 493
Db	241	PKHMGDSFAVEIKGRIHDFPQVVRITNIPFTYLEAVEAVLGKTSKIGAEETAI	300	Db 1261 VFLNSLGVVSVSTKIGFDSGGVRYVINEDLPELTQSROKNVTPNLPKEVLEIFQRKFQ 1320
Qy	301	WEDEESMKLQASMSMEDARATELGKEFPMELAKLIGOSWWDYSSSGNLVSNWLL	360	Qy 494 ----- 493
Db	301	WEDEESMKLQASMSMEDARATELGKEFPMELAKLIGOSWWDYSSSGNLVSNWLL	360	Db 1201 LAKSLCGVTAENKRKPSIFIDSSEPVRAFLRAYFVGDDITHPSKRLRLSTKSELLANQL 1260
Qy	361	RYAVERNELAPKPKPDESYRRLRTTYLGGVYKEPGRGLWNTITYDFRC	420	Qy 494 ----- 493
Db	361	RYAVERNELAPKPKPDESYRRLRTTYLGGVYKEPGRGLWNTITYDFRC	420	Db 1321 KNTEFEKELADSGKLDRKVKVLLDPLNLDIVLDRVKVNEKREVEGYVYDLSVEDNEN 1380
Qy	411	-----	410	Qy 494 ----- 493
Db	411	-----	410	Db 1381 FLVGFGLLYAHNSYGYMGYPKARWYSKECAESVTWGRHYIEMTIKEIEKFGFKVLA 1440
Qy	421	GKGIVNISDVKEGYDYLIGDGMQWVKKVHYEGKLININGIKCTPNHNKPVWVTENDRQ	480	Qy 543 D ----- 543
Db	421	GKGIVNISDVKEGYDYLIGDGMQWVKKVHYEGKLININGIKCTPNHNKPVWVTENDRQ	480	Db 1441 DSVTGDETEIIVRNGRTEFPVPIEKLFRDVYRIGEKEYCILEDVEALTLDRNGKLIKKV 1500
Qy	441	-----	410	Qy 544 ----- 543
Db	441	-----	410	Db 1501 PYVMRHRKKVYRWTNTSWYIDVTDHSLTVAEDGLKKEARPMETEGKSLIATKDDLG 1560
Qy	451	DIEYFDSRGKKKRISHQYRVEITIGENEKELLERILKIDFLGIRPSVKKGDTALKI	600	Qy 544 ----- 543
Db	451	DIEYFDSRGKKKRISHQYRVEITIGENEKELLERILKIDFLGIRPSVKKGDTALKI	600	Db 1561 VEYIKPHAEIISYNGVYDIEVEGTHRFANGILVHNTDRYATIPGKEPKETIKKAKE 1620
Qy	456	FLKYINSKLPGLELEVEFGYFLGFFVAKRYAVIDEEGRITTRGELEVRRDWSETAKT	625	Qy 566 FLKYINSKLPGLELEVEFGYFLGFFVAKRYAVIDEEGRITTRGELEVRRDWSETAKT 625
Db	456	FLKYINSKLPGLELEVEFGYFLGFFVAKRYAVIDEEGRITTRGELEVRRDWSETAKT	625	Db 1621 FLKYINSKLPGLELEVEFGYFLGFFVAKRYAVIDEEGRITTRGELEVRRDWSETAKT 1680
Qy	626	OAKVLEALKEDSVERAKETIVKDVVEIAYQVPLERKLVHIEQITKDLSEYKAIGPHVAI	685	Qy 626 OAKVLEALKEDSVERAKETIVKDVVEIAYQVPLERKLVHIEQITKDLSEYKAIGPHVAI 685
Db	626	OAKVLEALKEDSVERAKETIVKDVVEIAYQVPLERKLVHIEQITKDLSEYKAIGPHVAI	685	Db 1681 OAKVLEALKEDSVERAKETIVKDVVEIAYQVPLERKLVHIEQITKDLSEYKAIGPHVAI 1740
Qy	686	AKRLAANGIKYRPGTISYVURGSGKISDRVILUSYDPPKKHVKYDYYLQNVLPAVL	745	Qy 686 AKRLAANGIKYRPGTISYVURGSGKISDRVILUSYDPPKKHVKYDYYLQNVLPAVL 745
Db	686	AKRLAANGIKYRPGTISYVURGSGKISDRVILUSYDPPKKHVKYDYYLQNVLPAVL	745	Db 1741 AKRLAANGIKYRPGTISYVURGSGKISDRVILUSYDPPKKHVKYDYYLQNVLPAVL 1800
Qy	746	RLEAFGSKRKEKDLYQSSQVGLDAMKK	774	Qy 746 RLEAFGSKRKEKDLYQSSQVGLDAMKK
Db	746	RLEAFGSKRKEKDLYQSSQVGLDAMKK	774	Db 1801 RLEAFGSKRKEKDLYQSSQVGLDAMKK
Qy	721	VREMRKRNNSFYNLSFPEVSESYKGEVBDITLEGNPVYFANGILTHSLYPSIVTH	780	Qy 721 VREMRKRNNSFYNLSFPEVSESYKGEVBDITLEGNPVYFANGILTHSLYPSIVTH 780
Qy	421	VSPDTLREGCNYDAPIVGKFCDFPGFIPSITIGELITMRQELKKMKAIDIEKK	480	Qy 421 VSPDTLREGCNYDAPIVGKFCDFPGFIPSITIGELITMRQELKKMKAIDIEKK
Db	781	VSPDTLREGCNYDAPIVGKFCDFPGFIPSITIGELITMRQELKKMKAIDIEKK	480	Db 481 VSPDTLREGCNYDAPIVGKFCDFPGFIPSITIGELITMRQELKKMKAIDIEKK
Qy	481	MEDYRQAVKLLA-----	493	Qy 481 MEDYRQAVKLLA-----
Db	841	MEDYRQAVKLLA-----	493	Db 841 MEDYRQAVKLLA-----
Qy	494	-----	493	Qy 494 ----- 493
Db	494	-----	493	Db 494 ----- 493
Qy	901	VDNIFAPSLNKEKKSEIKVVALIRHKYKGEAEYEVLSGRKIHITRGSLSFTIRNGKI	960	Qy 901 VDNIFAPSLNKEKKSEIKVVALIRHKYKGEAEYEVLSGRKIHITRGSLSFTIRNGKI 960
RESULT	2			GN
DEOL_TheLLI				
TD	DEOL_TheLLI			
AC	STANDARD			
P30317;				
DT	01-APR-1993 (Rel. 25, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	DNA polymerase (EC 2.7.7.) (Vent DNA polymerase) [Contains: Endonuclease PI-T1II (EC 3.1.-) (Tli pol-1 intein) (IVPS2); DE Endonuclease PI-T1II (EC 3.1.-) (Tli pol-2 intein) (IVPS1)].			
DE	Endonuclease PI-T1II (EC 3.1.-) (Tli pol-2 intein) (IVPS1).			
POL.				

OS Thermococcus litoralis.
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Thermococcus.
 OX NCBI_TaxID=2265;
 RN [1]
 RF SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=930285; PubMed=608969;
 RA Perler F.B., Comb D.G., Jack W.E., Moran L.S., Qiang B., Hempstead S.K.,
 RA Kucera R.B., Bannier J., Slakko B.E., Nwankwo D.O., Hempstead S.K.,
 RT "Intervening sequences in an Archaea DNA Polymerase gene.;"
 RL Proc. Natl. Acad. Sci. U.S.A. 89:5577-5581(1992).
 RN [2]
 RP PROTEIN SPlicing.
 RX MEDLINE=9311708; PubMed=1475179;
 RA Hodges R.A., Perler F.B., Norren C.J., Jack W.E.; in an archaea DNA
 RT "Protein splicing removes intervening sequences in an archaea DNA
 RT Polymerase;"
 RL Nucleic Acids Res. 20:6153-6157(1992).
 CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
 EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
 -1- FUNCTION: INTEIN ENCODED ENDONUCLEASES ARE THOUGHT TO MEDIATE
 INTEIN MOBILITY BY SITE-SPECIFIC RECOMBINATION INITIATED BY
 ENDONUCLEASE CLEAVAGE AT THE "HOMING SITE" IN GENE THAT LACK THE
 INTEIN.
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 + [DNA](N).
 CC -1- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPlicing THAT INVOLVES
 A POST-TRANSLATIONAL EXCISION OF THE TWO INTERVENING REGION
 (INTEINS) FOLLOWED BY PEPTIDE LIGATION.
 CC -1- BIOTECHNOLOGY: Used in the PCR method because of its high
 thermostability and low error rate. Sold by New England Biolabs.
 CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
 CC ENDONUCLEASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to licensee@isb-sib.ch).

CC
 DR EMBL; M74198; AAA2100.1; -;
 DR EMBL; M74198; AAA2101.1; -;
 DR PIR; S42459; S42459.
 DR HSSP; P5689; 1TG0.
 DR REBASE; 2621; PI-TLII.
 DR REBASE; 2621; PI-TLII.
 DR InterPro; IPR002064; DNA_polymerase.
 DR InterPro; IPR003586; HintC.
 DR InterPro; IPR003587; HintN.
 DR InterPro; IPR002203; Intein.
 DR InterPro; IPR004042; Intein_endonuc.
 DR pfam; PF00136; DNA_Pol_B; 3.
 DR pfam; PF00104; DNA_Pol_B_exo; 1.
 DR PRINS; PR00379; INTEIN.
 DR SMART; SM00305; HintC; 2.
 DR SMART; SM00306; HintN; 2.
 DR SMART; SM00486; POLBC; 1.
 DR PROST; PS00116; DNA_Polymerase_B; FALSE_NEG.
 DR PROST; PS00810; PROTEIN_SPlicing; 2.
 KW Transferase; DNA-directed DNA Polymerase; DNA replication;
 KW DNA-binding; Hydrolase; Endonuclease; Autocatalytic cleavage;
 KW Protein_splicing; Intron homing;
 FT CHAIN 1 494 DNA POLYMERASE, 1ST PART.
 FT CHAIN 495 1032 ENDONUCLEASE PI-TLII.
 FT CHAIN 1033 DNA POLYMERASE, 2ND PART.
 FT CHAIN 1082 1471 ENDONUCLEASE PI-TLII.
 FT CHAIN 1472 1702 DNA POLYMERASE, 3RD PART.
 SQ SEQUENCE 1702 AA; 197293 MW; 21D6B98C75F3B20 CRC64;

Query Match 83.4%; Score 3356; DB 1; Length 1702;
 Best Local Similarity 42.7%; Pred. No. 3.6e-181; Gaps 2;
 Matches 726; Conservative 33; Mismatches 15; Index 928; Gaps 2;
 OY 1 MIFDFTYITKDGKPIRKFKEKPFKEKPIELDPHFRPYIYALKDSDAIBKAIKGERHG 60
 Db 1 MILDFTYITKDGKPIRKFKEKPFKEKPIELDPHFRPYIYALKDSDAIBKAIKGERHG 60
 OY 61 KIVRYDAVKKKLGROEVWKLIFEPHQDPAVLGRKREHPAVDIEYDIPFAKY 120
 Db 61 KTVRVIDAVKVRKKFGLGREVEWKLIFEPHQDPAVLGRKREHPAVDIEYDIPFAKY 120
 OY 121 LIDKGILPMCGDCEELKLMADIETFYHEDGPKGEKIMTSYADEEPEARYTWKNIDLPY 180
 Db 121 LIDKGILPMCGDCEELKLMADIETFYHEDGPKGEKIMTSYADEEPEARYTWKNIDLPY 180
 OY 181 VDVVSNEREMTKRFQVIVREKDPDLYTNGDNFDLPLYLIKRAEKLGVTLGLRKEHE 240
 Db 181 VDVVSNEREMTKRFQVIVREKDPDLYTNGDNFDLPLYLIKRAEKLGVTLGLRKEHE 240
 OY 301 WETEESMKLQAQYSMEDARATYLGKEFFMEALAKLIGQSVDWSRSSTGNLVEWLL 360
 Db 301 WETEESMKLQAQYSMEDARATYLGKEFFMEALAKLIGQSVDWSRSSTGNLVEWLL 360
 OY 361 RVAYERNEELAFLNPDPDEEYRRRLRTTYLGGYKEPERGIMENITYLDFRCLYPSITVHN 420
 Db 361 RVAYERNEELAFLNPDPDEEYRRRLRTTYLGGYKEPERGIMENITYLDFRCLYPSITVHN 420
 OY 421 VSPDTLERESCKKNYDVAPTYGKFCRDFPCTIPSITGELITMROBIKKMATAIDPIEK 480
 Db 421 VSPDTLERESCKKNYDVAPTYGKFCRDFPCTIPSITGELITMROBIKKMATAIDPIEK 480
 OY 494 ----- 493
 Db 541 VNNLFAFSNKKIKSESEVKVVKALIRHKYKGKAYEIQTLSSGRKINITAGHSFLTVNGEI 600
 OY 494 ----- 493
 Db 661 GMLRTLRRWMFGEENRRIRTFNRYLFHLEKLGLKLIPRGYEVTDWERLKKQLYEKL 720
 OY 494 ----- 493
 Db 601 KEVSGDGKIGEGLDLIVAPKIKLNEKGVSINIPELISDLSEETADIVMTISAKGRKNEFK 660
 OY 494 ----- 493
 Db 721 SVKYNGNKRKEVLVMENEIKDPISYFPOKELEEWKIGTLNGFRNCILKVDDEFGKLLGY 780
 OY 494 ----- 493
 Db 781 VSEGYAGAQKNTGGISYSVKLYNEDPVLLEMKAEGFGKVRDRKNCVSKKMAVL 840
 OY 494 ----- 493
 Db 841 VMKCLCGAIAENKRIPSVLTSPEPVRSFLEAYFTGDDIHPSKRFRFLSTKSELLAQ 900
 OY 494 ----- 493
 Db 901 VFLNLNSLGISSSVKIGFDGSGVYRVYINEDIQFPOTSPREKNTYYSNLIPKSLRLDVGKEFQ 960
 OY 494 ----- 493
 Db 961 KNMTFKKELVDSGKLNLREKAKLLEFFINGDIVLDRVSVKEKDYEYGVYDLSVEDNEN 1020
 OY 494 -----NSYYGMYGPKARWYSKECAESVTAWGRHYIEMTIKEEKEFGFKVLY 542

Db 419 VSPDTLNRENCKEYDVAPQVGHFRCKDFGPFPSLQLNLBEROKKRMKESKDPVEKK 478
 Qy 481 MLDYRORAVKLHANSYGYMGYPKARWYSECAESVTAWGRHYIEMTIKEEKGFKVL 540
 Db 479 LDYDQRRAIKLANSYGYGGYAKRWYSECAESVTAMGRQYDLYRRELESS-GFKVL 537
 Qy 541 YADTOGTYATPGEKPTIKKAKERFLKYIINSKPGLELEBEYEGYLRFGEVAKRYVI 600
 Db 538 YIDTDOGLYATIPGAKPNEKEKALFKEVEYINSKPGLELEYEGYARGFFVTKKYALI 597
 Qy 601 DEEGRITTRGIEWVRDWESELAKEOAKVIALEILKEDSVKAELIVKDVEELAKYQPL 650
 Db 598 DEEGKJLWTRGKLEIWRKDWSELAKEOAKVIALEILKHDGVNDBAVKTIVKEVTEKLSYEIPP 657
 Qy 661 EKLVHEQIKDSEKAIGHPHVAKRLAAGKIGVRPGDITSYIULRGSGKISDRVILL 720
 Db 658 EKLVHEQIUTPRLSEKAIGCPHVAKRLAAGKIGVRPGDITSYIULRGSGPISRAIAI 717
 Qy 721 SEYDPKKHHKDPPYIENQVLPAVIRILEAFGRKEDJKIQSOKVQGLDAWIK 773
 Db 718 EEFDPKKHKADEVYIENQVLPAPVERILRAFGYRKEEDLRYQTKQVGLGAWLK 770

RESULT 7
 DPOI_PYRFU STANDARD; PRT; 775 AA.
 AC P80061; P05584;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA polymerase (EC 2.7.7.7) (Pfu polymerase).
 GN PF0212.
 OS Pyrococcus furiosus, and
 OS Pyrococcus woesei.
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
 OX NCBI_TaxID=2261, 2262;
 RN [1] SEQUENCE FROM N.A., AND SEQUENCE OF 1-12.
 RC SPECIES=P. furiosus; STRAIN=VCL / DSM 3658;
 MEDLINE=91181200; Pubmed=8441634;
 RA Demori T., Ishino Y., Toh H., Asada F., Kato I.;
 RT "Organization and nucleotide sequence of the DNA polymerase gene from
 the archaon Pyrococcus furiosus.",
 RL Nucleic Acids Res. 21:259-265(1993).
 RN [2]
 RP SEQUENCE OF 396-418 FROM N.A.
 RC SPECIES=P. furiosus;
 MEDLINE=92107639; Pubmed=1762925;
 RA Mathur J., Adams M.W.W., Callen W.N., Cline J.M.;
 RT "The DNA polymerase gene from the hyperthermophilic marine
 archaeabacterium, Pyrococcus furiosus, shows sequence homology with
 alpha-like DNA polymerases.",
 RL Nucleic Acids Res. 19:6932-6932(1991).
 RN [3]
 RP SEQUENCE OF 396-502 FROM N.A.
 RC SPECIES=P. furiosus;
 RA Mathur E.J., Adams M.W.W., Callen W.N., Cline J.M.;
 RL Submitted (NOV-1991) to the EMBL/GenBank/DDBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.,
 RC SPECIES="P. woesei"; STRAIN=DSM 3773;
 MEDLINE=98032910; Pubmed=9758761;
 RA Dabrowski S., Kur J.;
 RT "Cloning and expression in Escherichia coli of the recombinant his-
 tagged DNA polymerases from Pyrococcus furiosus and Pyrococcus
 woesei.",
 RL Protein Expr. Purif. 14:131-138(1998).
 RN [5] SIMILARITY TO OTHER POLYMERASES.
 RP MEDLINE=92253366; Pubmed=1579479;
 RA Forterre P.;

RT not testify for a specific relationship between archaeabacteria and
 eukaryotes";
 RL Nucleic Acids Res. 20:1811-1811(1992).
 CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
 CC EXHIBITS 3', TO 5', EXONUCLEASE ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: N deoxyribonucleoside triphosphate = N diphosphate
 + [DNA](N).
 CC -1- SUBUNIT: MONOMER.
 CC BIOTECHNOLOGY: Because Pfu DNA polymerase exhibits the lowest
 error rate of any thermostable DNA polymerase studied, it is
 routinely used for PCR. It is sold by Promega.
 CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; D12983; BRA02362; 1;
 DR EMBL; U84155; AAB67984; 1;
 DR PIR; S35543; S35543;
 DR HSSP; P56889; 1TGO.
 DR Intertpro; IPR002064; DNA_pol_B.
 DR Pfam; PF00136; DNA_pol_B; 1.
 DR Pfam; PF0310; DNA_pol_B_exo; 1.
 DR PRINTS; PRO0106; DNAPOLB.
 DR SMART; SM00486; POLBC; 1.
 DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.
 KW Transferase; DNA-directed DNA polymerase; DNA replication;
 KW DNA-binding; Hydrolase; Endonuclease.
 SQ SEQUENCE 775 AA; 90113 MW; 566083D9F4DBD40E CRC64;

Query Match 78.3%; Score 3151; DB 1; Length 775;
 Best Local Similarity 74.6%; Pred. No. 4.3e-170; Mismatches 576; Conservative 100; Indels 2; Gaps 1;

Matches 576; Conservative 94; Mismatches 60; Indels 2; Gaps 1;

Qy 1 MIFDIDYITKDKGPITRIFKKENGERKEKIELDPHDFQPYIYALKDDSAIDSIAKIGERRH 60
 Db 1 MILDVDTYITERGKPVIRLFKRENGKFKIENDTRFPRYIYALLRDSKIEVKKTTGERHG 60
 Qy 61 KIVRVVDAVKKKFGLGRDVWVKLIFEPHPODPALRGKREHPAVIDYDIPFAKRY 120
 Db 61 KIVRIVDVEKVEKFKLGPITWVKLYLEHPDOPVTPIREKVREHPAVDIFYDIPFAKRY 120
 Qy 121 LIDKGILPMGSCDEEKLMMADIEFTYHEGDFGKELIMSYADEEAVRITWKDILP 180
 Db 121 LIDKGILPMGEEELKILARADIETLYHEGEGFKGPILIMSYADENEAKVITWKDILP 180
 Qy 181 VDVVSNERREMKFQVIREKOPDVLITYQNDNFDPILYKRAEKGVLGTTLGRKDEHPE 240
 Db 181 VEVVSSEREMKFLRFLIREKDPDLYVNGDSFDPYLAKRAEKGFLKLTIGRGS-E 238
 Qy 241 PKIHRMGDSFAVEVKGRIHDFLPVYRRTNLPTVTLAYEAVLIGTKTSKLGAEIAI 300
 Db 239 PKMQRIGDMTAVEVKGRHFLDLYVTRINLPTTLEAVYEAIFGKPKVYADEIAKA 298
 Qy 301 WETESEMKKLAQYSMEDAARAYELKEFPEMAELAKLGQSWSWYWSRSSTGNLWYLL 360
 Db 299 WESGENLERVAKYMSMEDAKAYELGKELPMEIQLSRLVQSPQWLWRSSTGNLWYLL 358
 Qy 361 RVAYERNELAPNPKPDEEEYRRRLRTYLGGYKEBERGLMENITLDFRLPYSITVN 420
 Db 359 RKAYERNEVAPNPKPSEEYQRRRSYTGFGVKEPEKGFLMENIVLDFRLPYSITVN 418
 Qy 421 VSPDTLEREGKNYDVAPVYKFCDFPGFIPSILGELITMROEIKKMKATDIEKK 480
 Db 419 VSPDTLNLEGKNDYJAPQVHFKCDPFGFIPSILGELLEEROKKMKETQDIEKI 478
 Qy 481 MLDYRORAVKLHANSYGYMGYPKARWYSECAESVTAMGRHYIEMTIKEEKGFKVL 540

"The DNA polymerase from the archaeabacterium Pyrococcus furiosus does

Db	719	VRYNGNKRREVIEVFNAVDVISMPEELKEWRIGCTRNGFRMGIFVUDIDFAKULLGYV	778	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
Oy	494	-----	-----	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation
Db	779	SEGSARKWKGNOTGGNSYTVRLYNENDEVLDMDMEHLAKKEFGKVKRGKNTVEIPKRMAYII	838	CC	the European Bioinformatics Institute. There are no restrictions on its
QY	494	-----	-----	CC	use by non-profit institutions as long as its content is in no way
Db	839	FESLQGTLAENKRKVPEVIFTSSKGVRWAFLEGFIGDGVHPSKRVLSTIKSELVNLGY	898	CC	modified and this statement is not removed. Usage by and for commercial
QY	494	-----	-----	CC	entities requires a licence agreement (See http://www.isb-sib.ch/announce/
Db	899	LLNSLGVSIAKLGVDGSGVIRVYNEELKFTEYRKKNVHSHVPKDILKETFGVKFOK	958	CC	or send an email to license@isb-sib.ch).
QY	494	-----	-----	493	-----
Db	959	NISYKKFRELVEENGKLDREKAKRIEWLNGDIVLDRVVEKREYDGVYDLSYDDEDENF	1018	CC	-----
QY	494	-----	-----	493	-----
Db	1019	LAGFGFLYAHNSYYGGYGYAKARWCKECAESTVAGREYIEFVRKELKEFGKVLYID	1078	DR	EMBL; AP00007; BAA31074_1; -
QY	544	TDGFYATIPGKPEEKKAKFLKYINSKLPGILETEYFGFYLRGGFYAKKAVIDE	603	DR	HSSP; P56889; ITG0.
Db	1079	TDGLYATIPGAKPEEKKAKFLFYINSKLPGILETEYFGFYLRGGFYAKKAVIDE	1138	DR	InterPro; IPR002064; DNA_Pol_B.
QY	604	GRITRGTLEYVRDMSSEIAETQAKYLEALKEDPSVAKETVWDVVEETAKYQVPLEK	653	DR	InterPro; IPR003586; HintC.
Db	1139	GRKTRTRGLEVRDMSSEIAETQAKYLEALKEDPSVAKETVWDVVEETAKYQVPLEK	1198	DR	InterPro; IPR003587; HintN.
QY	664	VHEQITKDSEYKAIGPHVAKRRAKQGKVRPGTISYVURGSGKISDRVILSEY	723	DR	InterPro; IPR0042203; IntCin.
Db	1199	VYEQITRPLRHEYKAIGPHVAKRRAQGKVRPGMVGYIVLRGDPISKRAILAEF	1258	DR	InterPro; IPR004042; Intein_endonuc.
QY	724	DPKKHYKDPDYYIENQVLPAVLRLAEGYRKEKEDLWRQKTKQGTLTAWL	1307	DR	Pfam; PF0016; DNA_Pol_B; 2.
Db	1259	DLRKHKYDAEYIENQVLPAVLRLAEGYRKEKEDLWRQKTKQGTLTAWL	1307	DR	Pfam; PF03104; DNA_Pol_B_exo; 1.
RESULT	9	DPOL_PYRHO	STANDARD;	PRT;	1235 AA.
ID	DPOL_PYRHO	STANDARD;	-----	PRT;	1235 AA.
AC	059610;	-----	-----	-----	-----
DT	15-DEC-1998	(Rel. 37, Created)	-----	-----	-----
DT	15-DEC-1998	(Rel. 37, Last sequence update)	-----	-----	-----
DT	16-OCT-2001	(Rel. 40, Last annotation update)	-----	-----	-----
DE	-----	DNA Polymerase (EC 2.7.7.7) [Contains: Pho pol intein].	-----	-----	-----
GN	-----	POL OR PHI947 OR PHB047.	-----	-----	-----
OS	Pyrococcus horikoshii.	-----	-----	-----	-----
OC	Archaea; Buryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.	-----	-----	-----	-----
NCBI_TaxID	53953;	-----	-----	-----	-----
RN	[1]	SEQUENCE FROM N.A.	-----	-----	-----
RP	STRAIN=OP3;	-----	-----	-----	-----
RC	MEDLINE=99344137;	PubMed=9679194;	-----	-----	-----
RA	Kawarabayasi Y., Savada M., Horikawa H., Haikawa Y., Hino Y., Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y., Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Onifuku Y., Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K., Masuchi Y., Shizuya H., Kikuchi H.;	-----	-----	-----	-----
RT	"Complete sequence and gene organization of the genome of a hyper-	-----	-----	-----	-----
RL	DNA Res. 5:55-76 (1998).	-----	-----	-----	-----
CC	-1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate	-----	-----	-----	-----
CC	+ [DNA] (N)	-----	-----	-----	-----
-1- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES	-----	-----	-----	-----	-----
CC	A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)	-----	-----	-----	491
CC	FOLLOWED BY PEPTIDE LIGATION (POTENTIAL)	-----	-----	-----	-----
CC	-1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.	-----	-----	-----	538
CC	-----	-----	-----	-----	492

Db	539 VKDLKLALSINRETKKSELKKVAKLIRRYSKGKVSILKSGRRKITSCHSLSFSVKNGKL
QY	492 -----
Db	599 VKVRGDELPGDLVWPGIKLPESKQVNLVELLKLBEEITSNIWMMIVPKGRKNEFK
QY	492 -----
Db	719 LKYGNNSRAYMVEFNSLRDVSLMPIEELKEWIGEPRPKIGFTIDVVDSEFAKLGLYYI
QY	492 -----
Db	659 GMLKTLWIFGEGEPRTRAGRLKHLERIGYVKLRRGGEVLDWESLKKRYKLYETLKN
QY	492 -----
Db	779 SSGDVEKDKVFKHSDKDNYLEDIKAELAKLAEKLFKGKVRRGRGYIEVSGKISHAFLVLAEGRR
QY	494 -----
Db	839 IPEFIFTSPMDIKVAFKLKGNGNAELTFSTKSBLVNLNQILLNSIGVSDIKEHEKGV
QY	494 -----
Db	899 YRVINKKESSENGDIDVDSVESIEVEKYGVYDLSVENEFLVFGFLYAHNSYGGY
QY	501 GYPKARWYSKESCAESVTAVGRHYEMTKEIEERFGPKLYATDGFYKTPGKPK-ETI
Db	959 GYAKARWYKCEKACSVTAANGQYIDLVRELEAR-GFKVLYIDFDGLYATIPGVKDWEV
QY	560 KKKAKEFLKYINSKLPGLLELEFGYLGFFVAKKRYAVIDEGRITRGLEVVRDGS
Db	1018 KRRALFAYVYDINSKLPGLLEFGYFVAKKRYAVIDEGRITRGLEVVRDGS
QY	620 EIAKETQARVLEATLKGHVEAVKIVDQVTEKINTYEVPPKEKIVYQITRPINEKAI
Db	1078 EIAKETQARVLEATLKGHVEAVKIVDQVTEKINTYEVPPKEKIVYQITRPINEKAI
QY	680 GPHVIAKRLAAGKIKVRGTTISYIVLGGSKSDRVYLSEDPKKHRYDDYIENQ
Db	1138 GPHVAVAKRLMARGIKVKGWGMVIGIVLKGDPKISKRAISIEEDPDKRKYDAEYIENQ
QY	740 VLVAVLRLTEAEGYRKEDQYQSSQKVQGLDAWIK
Db	1198 VLPVAVERILKAFGKREDLWRQKTKQVGLGAWIK
CC	1231
RESULT	10
COL_THERM	STANDARD; PRT; 1523 AA.
DPOL_THERM	AC 4918; DT 15-DEC-1998 (Rel. 37, Created) DT 15-DEC-1998 (Rel. 37, Last sequence update) DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE	DNA Polymerase (EC 2.7.7.7) (Pol I tfu) [Contains: Endonuclease PI-Tfu (EC 3.1.-.) (tfu pol-1 intein); Endonuclease PI-Tfui (tfu pol-2 intein)].
GN	POLE
OS	Thermococcus fumicolaens.
OC	Archaea; Eurarchaeota; Thermococcales; Thermococcaceae; Thermococcus.
OX	NEBI-TAXID=46540;
RN	[1] SEQUENCE FROM N.A.
RC	STRAIN-ST57;
RA	Clibon M., Quereilou J.; Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RL	[2] CHARACTORIZATION OF INTEINS. STRAIN-ST57; MEDLINE=20112788; PubMed=10644683; SAVES I., Orenne V., Dietrich J., Masson J.-M.; Inteins of Thermococcus fumicolaens DNA polymerase are endonucleases
RT	with distinct enzymatic behaviors. ";
RL	J. Biol. Chem. 275:2335-2341(2000).
CC	-!- FUNCTION: PI-Tfu recognizes and cleaves a minimal sequence of 16 base pairs (bp) on supercoiled DNA with either Mn(2+) or Mg(2+) as cofactor. It cleaves linear DNA only with Mn(2+) and requires a 19-bp minimal recognition sequence. The optimal temperature for activity is 70 degrees Celsius.
CC	-!- FUNCTION: PI-Tfui is a highly active homing endonuclease using Mg(2+) as cofactor. Its minimal recognition and cleavage site is 21 bp long either on linear or circular DNA substrates. Its endonuclease activity is strongly inhibited by the 3' digestion product, which remains bound to the enzyme after the cleavage reaction. The optimal temperature for activity is 70 degrees Celsius.
CC	-!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate + (DNA)(N).
CC	-!- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPlicing THAT INVOLVES A POST-TRANSLATIONAL EXCISION OF THE TWO INTERVENING REGION (INTEIN) FOLLOWED BY PEPTIDE LIGATION.
CC	-!- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
CC	-!- SIMILARITY: IN THE INTEIN SECTION; BELONGS TO THE HOMING ENDONUCLEASE FAMILY.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	EMBL: Z69882; CAA93738.1; -.
DR	HSSP: P56889; 1TGO.
DR	REBASE; 4500; PI-Tfu.
DR	InterPro: IPR002064; DNA_POL_B.
DR	InterPro: IPR003586; HintC.
DR	InterPro: IPR003587; HintN.
DR	InterPro: IPR002203; Intein.
DR	InterPro: IPR004042; Intein_endonuc.
DR	Pam: PFO013; DNA_Pol_B; 3.
DR	Pfam: PF03104; DNA_Pol_B_exo; 1.
DR	PRINTS: PR03179; INTEIN.
DR	SMART: SM00305; HintC; 2.
DR	SMART: SM00306; HintN; 2.
DR	PROSITE: PS00116; DNA_POLYMERASE_B; FALSE_NEG.
DR	PROSITE: PS0081; PROTEIN_SPlicing; 2.
KW	Transfere: DNA-directed DNA polymerase; DNA replication; DNA-binding; Hydrolase; Endonuclease; Autocatalytic cleavage; Protein Splicing; Intrtron homing; Magnesium; Manganese.
FT	CHAIN 1 DNA POLYMERASE, 1ST PART.
FT	CHAIN 407 766 ENDONUCLEASE PI-TFU.
FT	CHAIN 767 900 DNA POLYMERASE, 2ND PART.
FT	CHAIN 901 1282 ENDONUCLEASE PI-TFUI.
FT	SEQUENCE 1283 1523 DNA POLYMERASE, 3RD PART.
SQ	SEQUENCE 1523 AA; 1/5917 MW; 7A2AC8236BF2E5F5 CRC64;
Query	69.8%; Score 2809; DB 1; Length 1523;
Match	Best Local Similarity 39.0%; Pre. No. 1.8e-150; Mismatches 91; Indels 752; Gaps 4;
QY	1 MIFDFTYITRKDGPKIIRIKRKENGKFLDLPHCPYVALLKODSAIDEKAJKGERHG
Db	1 MILDFTYITEDGDPWIRVKFKENGKFLKIEVDPFPYVALLKODSAIDEVKKITSRHG
QY	61 KIVRVDAVYKVKKFLGRVEWVWKLIFERQDYLPAURGKIREPAVIDYKXIPFARRY
Db	61 TTVRVVRAGKVKKKLFLGRPEWVKKLYFTHQDPAIRKIREPAVWDYIYDPAFKRY
QY	121 LIDKGGLIPMEGDEEKFMAFDIETFYHEQDEFGKGKELMISYADEEVARVITWNKIDP

Db	121 LIDKGKLIPEGDEEJKMLAFDIEITYHEGEFAEOPILMSYADEBEARWITKKIDLPY	180	Qy	544 -----	Qy	544 -----	Db	1258 RKEEITDYGVYDIEVEGTHFFANGILVHNTDSFFATIFGADAEVKKKAREFLNIN	1317
Qy	181 VDVVNEREMRIKREVIOVREKDPDLYTNGDNFOLPLYKRAEKLGVTLLGRKEHE	240	Qy	544 -----	Qy	544 -----	Db	1259 RKLPGILLELEYEGFYLRGFFVAKKRYAVILEBEGRTTRGLEVWDRDWSLAKETOAVL	631
Db	181 DVVSTEKEMKIKRFLVKWERKDPLVITYNGDNFDFAYLKKRSEKLGVKFILGRGS-E	238	Qy	572 SKLPGILLELEYEGFYLRGFFVAKKRYAVILEBEGRTTRGLEVWDRDWSLAKETOAVL	631	Qy	572 SKLPGILLELEYEGFYLRGFFVAKKRYAVILEBEGRTTRGLEVWDRDWSLAKETOAVL	631	
Qy	241 PKTHRMGDSFAVEIKGRIRHDFPVVRITINLPYTLAEVAVLGTKSKLGAEEIAI	300	Qy	572 SKLPGILLELEYEGFYLRGFFVAKKRYAVILEBEGRTTRGLEVWDRDWSLAKETOAVL	631	Qy	572 SKLPGILLELEYEGFYLRGFFVAKKRYAVILEBEGRTTRGLEVWDRDWSLAKETOAVL	631	
Db	239 PKIORMGDRFAVEVKGRIHDFPVTRHTNLPYTLAEVAVLFGOPKEKVYEEIAQ	298	Qy	632 ALKEDSVKAEVTKDVEELAKQVPLKWLHQITKDLSEKAIGHVAIKLAA	691	Qy	632 ALKEDSVKAEVTKDVEELAKQVPLKWLHQITKDLSEKAIGHVAIKLAA	691	
Qy	301 WEETEESMKKLAQYSMEDARATYELCKEFFMEEAELAKLIGQSVDVSRSSTGNGLVEWLL	360	Qy	632 ALKEDSVKAEVTKDVEELAKQVPLKWLHQITKDLSEKAIGHVAIKLAA	691	Qy	632 ALKEDSVKAEVTKDVEELAKQVPLKWLHQITKDLSEKAIGHVAIKLAA	691	
Db	299 WETGGLERARYMSMEDAKTYELGREFFMEAQLSRLQGQSFDNSRSSTGNGLVEWLL	358	Qy	632 ALKEDSVKAEVTKDVEELAKQVPLKWLHQITKDLSEKAIGHVAIKLAA	691	Qy	632 ALKEDSVKAEVTKDVEELAKQVPLKWLHQITKDLSEKAIGHVAIKLAA	691	
Qy	361 RVAYERNELAPNPKDDEEYRRRLRTYLGGVYKEPERGLMENTITYLDFRC-----	410	Qy	692 KGIKURPGTISYIVRGSGKISDRVLLSEDPKKHPPDYYENQVPAVRLAEF	751	Qy	692 KGIKURPGTISYIVRGSGKISDRVLLSEDPKKHPPDYYENQVPAVRLAEF	751	
Db	359 RKAVERNELAPNPKSRELER-RGGYAGGYKEPERGLMENTIAYLDFRCHPADTKV	417	Qy	692 KGIKURPGTISYIVRGSGKISDRVLLSEDPKKHPPDYYENQVPAVRLAEF	751	Qy	692 KGIKURPGTISYIVRGSGKISDRVLLSEDPKKHPPDYYENQVPAVRLAEF	751	
Qy	411 -----	410	Qy	752 GYRKEDLYKQSSKQVGLDANIK	773	Qy	752 GYRKEDLYKQSSKQVGLDANIK	773	
Db	418 GKGVVNISEVREGDYVLGIDQWQKVQRWVNDYEGELVNINGLCTPNHKLPPVRRTERQ	477	Qy	1498 GYKEDLYQKTRQVGLGAWLK	1519	Qy	1498 GYKEDLYQKTRQVGLGAWLK	1519	
Qy	411 -----	410	Db	RESULT 11		Db	RESULT 11		
Db,	478 TAIRSLAKSFLTKKVKGKLTTPLEFEKGKIERDPEPELKGELAGITLARGTLLRK	537	DPOL_PYRKO	IDPOL_PYRKO	STANDARD;	Db	IDPOL_PYRKO	STANDARD;	
Qy	411 -----	410	Qy	411 -----	410	Qy	411 -----	410	
Db	538 DVEYFDSSRGKKRVSQHQRVREITVGAQEEDFQRRIVYIFERLFGVTPSVRKKNIAITF	597	DT	16-OCT-2001 (Rel. 40, Created)	DT	Db	1318 PKLPGILLELEYEGFYLRGFFVAKKRYAVILEBEGRTTRGLEVWDRDWSLAKETOAVL	1377	
Qy	411 -----	410	DT	16-OCT-2001 (Rel. 40, Last sequence update)	DT	Qy	632 ALKEDSVKAEVTKDVEELAKQVPLKWLHQITKDLSEKAIGHVAIKLAA	691	
Db	598 KVAKKEVYLVRREIMDGIEMLHAPSVLRGFFEGDSVNUVKVRKTVVVNQGTNNENKIEVVS	657	DE	DNA Polymerase (EC 2.7.7.7) (Contains: Endonuclease PI-PK01 (EC 3.1.-.) (Pko pol-1 intein) (IVS-A); Endonuclease PI-PK01 (EC 3.1.-.) (Pko pol-1 intein) (IVS-B)).	DE	Db	1438 RGKURPGTIVSYIVKGSGRIGDRPDEFDPKXHRYDAEXYIENQVPAVERIKAF	1497	
Qy	411 -----	410	DN	POL.	DN	Qy	752 GYRKEDLYKQSSKQVGLDANIK	773	
Db	658 KLLNKGIPHRRYTYDTEERKTMTHILETAGRGLLIFQTYVGFISTEKNALEAIR	717	OS	Pyrococcus kodakaraensis.	OS	Qy	752 GYRKEDLYKQSSKQVGLDANIK	773	
Qy	411 -----	410	OC	Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Thermococcus.	OC	Qy	752 GYRKEDLYKQSSKQVGLDANIK	773	
Db	718 NRERVRLENNAFTYTLADFTAKTEVYKGKVYDITLEGTPYFANGILTHNSLYPSIISHN	777	OX	NCB1_TaxID=6914;	OX	Qy	752 GYRKEDLYKQSSKQVGLDANIK	773	
Qy	421 VSPDPLEREGCKKYYVAPTYGKCKDFGFFIPSTIGELITMROBEIKKKKATIDPIEK	480	RN	[1]	RN	Qy	752 GYRKEDLYKQSSKQVGLDANIK	773	
Db	778 VSPDPLNREGCGEYDEAPOVGHRECKDFPGFFPLSGLDILDEORVKHKMATAVDPIEK	837	RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.	RP	Qy	752 GYRKEDLYKQSSKQVGLDANIK	773	
Qy	481 MLIDYIQORAVKLHANSYYGKPKARWYSECAESTANGRHYIEMTKEIEERGFV	540	RC	STRAT=KOD1;	RC	Qy	752 GYRKEDLYKQSSKQVGLDANIK	773	
Db	838 LLDRYRORAKILANSFYGGYAKARWYCKECAESVTANGRQYLTETMREIEERGFV	897	RC	MEDLINE=98027; PubMed=936136;	RC	Qy	752 GYRKEDLYKQSSKQVGLDANIK	773	
Qy	541 YAD-----	543	RT	Takagi M., Nishioka M., Kakinara H., Kitabayashi M., Inoue H., Kawakami B., Oku M., Imanaka T.; "Characterization of DNA polymerase from Pyrococcus sp. strain KOD1 and its application to PCR"; Appl. Environ. Microbiol. 63:4504-4510(1997).	RT	Qy	752 GYRKEDLYKQSSKQVGLDANIK	773	
Db	898 YADSVTGDTEVTRRNGRRIEFPVIEKLFERVDHRVGEKEVCKVULGVEALTLDNRGRVLW	957	RT	[2]	RT	Qy	752 GYRKEDLYKQSSKQVGLDANIK	773	
Qy	544 -----	543	RL	CHARACTERIZATION OF INTEINS.	RL	Qy	752 GYRKEDLYKQSSKQVGLDANIK	773	
Db	958 KPYVYMRHKTDKRYRVWFTNSWLDVTEHDHSILYNTSKVPGKPLKERLVEVKEEL	1017	CC	CC	CC	Qy	752 GYRKEDLYKQSSKQVGLDANIK	773	
Qy	544 -----	543	CC	CC	CC	Qy	752 GYRKEDLYKQSSKQVGLDANIK	773	
Db	1018 GGKVSLITPNRPIARTIKANPIAVKLWELIGLIVGDNNGGOSWAKYVGLSCLGLDKA	1077	CC	CC	CC	Qy	752 GYRKEDLYKQSSKQVGLDANIK	773	
Qy	544 -----	543	CC	CC	CC	Qy	752 GYRKEDLYKQSSKQVGLDANIK	773	
Db	1078 EIERKVLNPLREASIVSINYDKSKKGDSVLSKMLAGFMVKYFDENGKNAIPSEMFNLP	1137	CC	CC	CC	Qy	752 GYRKEDLYKQSSKQVGLDANIK	773	
Qy	544 -----	543	CC	CC	CC	Qy	752 GYRKEDLYKQSSKQVGLDANIK	773	
Db	1138 REYIEAFLRLGLFSDAGTVSLRGGIPEIRLTSVNRLELSDAVRKLWLVGSNSLFTETKPN	1197	CC	CC	CC	Qy	752 GYRKEDLYKQSSKQVGLDANIK	773	
Qy	544 -----	543	CC	CC	CC	Qy	752 GYRKEDLYKQSSKQVGLDANIK	773	
Db	1198 RYKEKESGTHSIHVRIKHNRFADRGFLIDRKSTKLSLENGLGHNTKRAYKBFDFLV	1257	CC	CC	CC	Qy	752 GYRKEDLYKQSSKQVGLDANIK	773	

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC
or send an email to license@ib-sib.ch).

CC
DR EMBL; D29671; BAA06142.2; -.
DR HSPB; P56689; IntC.
DR REBASE; 3792; PI-PK1.
DR REBASE; 3793; PI-PK1.
DR InterPro; IPR00264; DNA_Pol_B.
DR InterPro; IPR002586; HintC.
DR InterPro; IPR002203; Intein.
DR InterPro; IPR004042; Intein_endonuc.
DR Pfam; PF00156; DNA_Pol_B; 3.
DR Pfam; PF03104; DNA_Pol_B_exo; 1.
DR PRINTS; PR03379; INTEIN.
DR SMART; SM00305; HintC; 2.
DR SMART; SM00306; HintN; 2.
DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.
DR PROSITE; PS00881; PROTEIN_SPLICING; 2.
DR transferase; DNA-directed DNA polymerase; DNA replication;
DNA-binding; Hydrolase; Endonuclease; Autocatalytic cleavage;
Protein splicing; Intron homing.
DR CHAIN 1 405 DNA_POLYMERASE_1ST_PART.
FT CHAIN 407 766 ENDONUCLEASE_PI-PRO_I.
FT CHAIN 852 851 DNA_POLYMERASE_2ND_PART.
FT CHAIN 1389 1671 ENDONUCLEASE_PI-PRO_IT.
FT SEQUENCE 1671 AA; 193490 MW; 4A17FC1B120EE455 CRC64;

Query Match 69.1%; Score 2782; DB 1; Length 1671;
Best Local Similarity 36.0%; Pred. No. 6_9e-19; Mismatches 82; Indels 900; Gaps 4;
Matches 602; Conservative 82; Mismatches 86; Indels 900; Gaps 4;

Qy 1 MIFDTDTYTKDCKPTRIFKKGKNGEKFKEKELDPHQPTYALLKKDSADEIKAICKEHRG 60
Db 1 MILDIDITTEDGKPKVIRKKGKNGEKFKEKELDPHQPTYALLKKDSADEIKAICKEHRG 60
Qy 61 KIVVWVDAVAKVKKKGKELGRDVEWKLIFHEPQDPAVLRKGKIREHAPVIDEYDIPAKRY 120
Db 61 TVVTVKRVKEVOKKFLGRPVVEWKLYTHPQDPAVLRKDTRKIREHAPVIDEYDIPAKRY 120
Qy 121 LIDKGKLFPMGDEEKLIMAFDTEFREGDEGGKELMISADEEVARVITWKNDLPY 180
Db 121 LIDKGKLFPMGDEEKLIMAFDTEFREGDEGGKELMISADEEVARVITWKNDLPY 180
Qy 181 VDVVSNSNREMIRKFVQTVREKOPDVLYTYNGDNFDIPLYLTIRKRAEKLGVTLLIGRDKEHPE 240
Db 181 VDVVSNSNREMIRKFVQTVREKOPDVLYTYNGDNFDIPLYLTIRKRAEKLGVTLLIGRDKEHPE 240
Qy 241 PRKHMDSFAVEIKGTHDFPFPVVRRTINPPTYLEAVEAIVGKTKSKIGAEETAI 300
Db 241 PRKHMDSFAVEIKGTHDFPFPVVRRTINPPTYLEAVEAIVGKTKSKIGAEETAI 300
Qy 299 WETGGENIERYAVSMSMEDAKVYTLGKREFLPMHEAQLSLRIGLQSLWDYRSSTGVLNVEWLL 358
Db 299 WETGGENIERYAVSMSMEDAKVYTLGKREFLPMHEAQLSLRIGLQSLWDYRSSTGVLNVEWLL 358
Qy 301 WETEESMKLAKQSMSMEDAREAYLGEFFPMHEAQLSLRIGLQSLWDYRSSTGVLNVEWLL 360
Db 301 WETEESMKLAKQSMSMEDAREAYLGEFFPMHEAQLSLRIGLQSLWDYRSSTGVLNVEWLL 360
Qy 361 RYAYERNELAPNRPDEEREYRRRLRTYLGGYKKEPERGLWENITYDFRC----- 410
Db 361 RYAYERNELAPNRPDEEREYRRRLRTYLGGYKKEPERGLWENITYDFRC----- 410
Qy 359 RKAERNELAPNRPDEELARR-RQSEGGIVKEPERGLWENIVLDFRCPADKWK 417
Db 359 RKAERNELAPNRPDEELARR-RQSEGGIVKEPERGLWENIVLDFRCPADKWK 417
Qy 411 ----- 410
Db 411 ----- 410
Qy 418 GKGGINSEVQEGDYVLGIDGWQVRKVWEVDYKGELVNINGLKCCTPNHKLPVVTNERQ 477
Db 418 GKGGINSEVQEGDYVLGIDGWQVRKVWEVDYKGELVNINGLKCCTPNHKLPVVTNERQ 477
Qy 411 ----- 410
Db 411 ----- 410
Qy 538 DVEFDSRKRRKISHQYRVEITIGKDEEFDRDITYIFERLFGTPSISEKKGNAVT 597
Db 538 DVEFDSRKRRKISHQYRVEITIGKDEEFDRDITYIFERLFGTPSISEKKGNAVT 597

Qy 411 ----- 410
Db 598 KVAKKNVYLKVKEIMDNIESLHAPSVLRGFFEGDSVNRVRSIVATQTKNEWKIKLV 657
Qy 411 ----- 410
Db 658 KLLSQLGIPHQTYIQYQENGKDRSRVILETGKDGILILFQTLIGFISERKNALLNKAIS 717
Qy 411 ----- 420
Db 718 QREMNNTENNGTYRLSERNSVTEYEVYKVDLTLEGTPYFANGILTHNSLYPSITITHN 777
Qy 421 VSPDTLEREGCKNYDVAPIVGKFKCQDFPGFIPSILGELITMRQETKKKMATDIEK 480
Db 778 VSPDTLNRGCCKEYDVAPQVGHRFCQKQFGFPLSLGDLEROKKMKATDIEK 837
Qy 494 ----- 493
Db 838 LDYRQRAKILANSILPEWLPUERGEVHVFRIGELIDRMEENAGKVKREGETEVLE 897
Qy 494 ----- 493
Db 898 VSGLEVPSNRRNKAELKRVKALRHDYSGSKVYTRLKSGRKITSGHSLSFSVRNGEL 957
Qy 494 ----- 493
Db 958 VEVTDGELKPGDLVAVPRLPEPNHVLNLVELLGTPPEETLDIVMTIPVKGKKNEF 1017
Qy 494 ----- 493
Db 1018 GMRLTRWIFGEEKRPRRTARRVIRHLLEDGVRLKKIGEVLDWDSLKNYRRLYEALVEN 1077
Qy 494 ----- 493
Db 1078 VVYGNKRELYVEFNSTRDAVGIMPLKELKEWKGITLNGFRMRKLEEVDESLAKLIGYY 1137
Qy 494 ----- 493
Db 1138 SEGYARKQRNPKNQWSYSVLYNEDPVEVLLDDMERLASRFGKVRGRGRNYVLPKKIGYL 1197
Qy 494 ----- 493
Db 1198 FENMCGVLAENKRIPFVFVTPSKGVRLAFLBEGFYFIGDDVHVNKLRLSTKSELLANOLV 1257
Qy 494 ----- 493
Db 1258 LLNSVQVSAVKGHDGSGVYRVVYINBELLFWKLKKKNAYVSHVTKEVLSVEFGVFKQ 1317
Qy 494 ----- 493
Db 1318 NVSPQTRKVMVEDGRDLPKPEKAQRLSLWILGEVDVLDLDRVESVVEDYDGYYVLSVEDNENF 1377
Qy 494 ----- 493
Db 1378 LVGGFLVYAHNSYYGGYKARWYKCAESVTANGREYVITMTEKEIERFGFKLYAD 543
Qy 544 TDGFYATIPGKPKETKKKPKFLKVKINSKPLGLEEYEGTFLGRFPAVKRYAVIDE 603
Db 1438 TDGFYATIPGKPKETKKKPKFLKVKINSKPLGLEEYEGTFLGRFPAVKRYAVIDE 1497
Qy 604 GRITTRSLERVVRDWSPIAKETOKYLEALALKEDSVKAEVTKDVVEETAKYQPLEK 663
Db 1498 GKITRGLETIVVRDWSPIAKETOKYLEALALKEDSVKAEVTKDVVEETAKYQPLEK 1557
Qy 664 VTHEQITKDLSEBYKAQGPHVIAKRLAAGKIGKVRPCTISIVLGSKLSDRVILSEY 723
Qy 1558 VHEQITRDLKDKYKATGPHVIAKRLAAGKIGKVRPCTISIVLGSKLSDRVILSEY 723
Db 724 DPKKHYDPPYIENQVLPAYLRLAEGYKVRPCTISIVLGSKLSDRVILSEY 773
Qy 1618 DPTKHYDAEVYIENQVLPAYLRLAEGYKVRPCTISIVLGSKLSDRVILSEY 1667

RESULT 12 QY 1 MFDIDYITKDGKPIRKENGKERKIELDPHPHQPYIYALLKDDSADETKAIGERHG 60

DPOL_THEG8 STANDARD; PRT; 1699 AA. Db 1 MLDTYTYTEGKPVIRVKKENGEKIEVDRNFEYFYALKDSAIEVKKITAHRG 60

ID DPOL_THEG8 Q9H8A; 16-OCT-2001 (Rel. 40, Created) QY 61 KIVRVVDAVVKKKFGLRDEVWKLIFERPHDPAVLRGKREHPAVPIDYEDIFAKRY 120

AC 09H8A; DT 16-OCT-2001 (Rel. 40, Last sequence update) QY 61 TIVVKVRAEKKKFGLRDEVWKLIFERPHDPAVLRGKREHPAVPIDYEDIFAKRY 120

DT 16-OCT-2001 (Rel. 40, Last annotation update) DE DNA Polymerase (EC 2.7.7.7) (Contains: Endonuclease PI-Tspg8II

DE (EC 3.1.-) (Tsp-GEB pol-1 intein); Endonuclease PI-Tspg8II

DE (EC 3.1.-) (Tsp-GEB pol-2 intein)). DE POL OR POL-1.

GN OS Thermococcus sp. (strain GEB). QY 181 VDVVSMERMKRKFQIVREKOPDVITYNQDNFDIPLYIKRAEKLVGTVLGRDKEPE 240

Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Thermococcus. OC NCBIX-TAXID=105583; QY 181 VDVVSMERMKRKFQIVREKOPDVITYNQDNFDIPLYIKRAEKLVGTVLGRDKEPE 240

RN [1] RT SEQUENCE FROM N.A. RA "Thermococcales taxonomy and phylogeny based on the comparative use of 16S rRNA, 16S-23S rRNA intergenic spacer and family B DNA polymerase genes"; RT Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases. CC - FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE EXHIBITS 3', TO 5', EXONUCLEASE ACTIVITY (BY SIMILARITY).

CC - FUNCTION: PI-TSPGE8I AND PI-TSPGE8II ARE ENDONUCLEASES (POTENTIAL).

CC - CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate + (DNA)(N).

CC - PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN) FOLLOWED BY PEPTIDE LIGATION.

CC - SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.

CC - SIMILARITY: IN THE INTEIN SECTION; BELONGS TO THE HOMING ENDONUCLEASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to licensee@isb-sib.ch).

CC EMBL; AJ25033; CAC12850; 1. DR 361 RVAYERNELAPNKPDEEYRRRLRTYLGYYKEBRLGLENITYLDFRCIPLSPVTHN 420

DR InterPro; IPR002064; DNA_Pol_B. DR 359 RKAVERNELAPNKPDEEYRRRLRTYLGYYKEBRLGLENITYLDFRCIPLSPVTHN 417

DR InterPro; IPR003586; HintN. DR 421 VSPDTLEREGKNDVAPVQYKFCDFPGFIPSILGELTMRQETKKKAKATDIEKK 480

DR InterPro; IPR003587; HintN. DR 418 VSPDTLEREGKNDVAPVQYKFCDFPGFIPSILGELTMRQETKKKAKATDIEKK 477

DR InterPro; IPR002203; Intein. DR 478 LDLYRORAKIILANSLPDEWLPLLVNGRLKLVIGDFVNTMKKGQPLENGTELEVS 537

DR InterPro; IPR004042; Intein_endonuc. DR 598 VTGGEIKPGGFIAVPRVNLPERHERINLIEILSLPPETSDIVLTPVKGRKNEFKGM 657

DR Pfam; PF00136; DNA_Pol_B; 3. DR 494 ----- QY -----

DR InterPro; IPR003104; DNA_Pol_B_exo; 1. DR 658 LRTLWIFEEQQRPTARRVLEHLQKLGYKLMKRAYETVNEAKLRNRYKLYEVAERVK 717

DR PRINTS; PRO0106; DNAPOLB. DR 493 -----

DR PRINTS; PRO0379; INTEIN. DR 493 -----

DR SMART; SM00305; HintN; 2. DR 494 -----

DR SMART; SM00306; HintN; 2. DR 493 -----

DR SMART; SM00486; PolB; 1. DR 493 -----

DR PROSITE; PS00831; PROTEIN_SPLICING; 2. DR 493 -----

DR PROSITE; PS00116; DNA_POLYMERASE_B; FALSE_NEG. DR 493 -----

DR TRANSFERASE; DNA-directed DNA polymerase; DNA replication; DNA-binding; Hydrolase; Nuclease; Exonuclease; Multifunctional enzyme; Protein_splicing; Autocatalytic cleavage; Endonuclease; Intron homing. DR 493 -----

DR TRANSFERASE; DNA-directed DNA polymerase; DNA replication; DNA-binding; Hydrolase; Nuclease; Exonuclease; Multifunctional enzyme; Protein_splicing; Autocatalytic cleavage; Endonuclease; Intron homing. DR 493 -----

DR CHAIN 1 491 DNA POLYMERASE, 1ST PART. DR 493 -----

FT CHAIN 492 1026 ENDONUCLEASE PI-TSPGE8II. DR 493 -----

FT CHAIN 1027 1075 DNA POLYMERASE, 2ND PART. DR 493 -----

FT CHAIN 1076 1464 ENDONUCLEASE PI-TSPGE8II. DR 493 -----

FT CHAIN 1465 1599 DNA POLYMERASE, 3RD PART. DR 493 -----

SQ SEQUENCE 1699 AA; 197323 MN; F389B4351F0B12D3 CRC64. DR 493 -----

Query Match 68.8%; Score 2769.5; DB 1; Length 1699; QY 494 -----

Best Local Similarity 35.5%; Pred. No. 3.5e-148; Matches 602; Conservative 85; Mismatches 83; Indels 927; Gaps 4; QY 544 -----

1077

Db	1078	AGNTEVITIRNGKVEFVPIEKLQFRVDYRIGEKEYCALEGVEALTLDRNRSGLWWRKVPI	1137	CC
QY	544	-----	-----	CC
Db	1138	MRHKTKKKIYKWFNTNSWLDVTEDHSLIGLNTSKVKSEKPLKERLVEVKPRLGEKVK	1197	CC
QY	544	-----	-----	543
Db	1198	SLITNRAIARSIKANPIAVRLWELIGLVLGDGNNGGHSKWAKEYVGLSGCLDKATEEK	1257	CC
QY	544	-----	-----	543
Db	1258	VLRPLKEAGITISNYGGSKKKGDVSLSKWLAGFMVKYKFQDENKNKRIPSMFNLPREYIE	1317	CC
QY	544	-----	-----	543
Db	1318	AFLRLGLFSADGTIVSLRRGPIPEIRLTSVNRLSNEVRKLINLVLGVNSNSMFTTPNKG	1377	CC
QY	544	-----	-----	543
Db	1378	ESGTRSIHVRIKKNHFKRKGFLDRKATKLSDNLRHNTKKMAYRIDFDLWPKIEE	1437	CC
QY	544	-----	-----	543
Db	1438	INYDRVYDIEVEGTHRPFANGILVHNTGDFATPGADAEVTKKAMEFVYKLNKPG	576	CC
QY	544	-----	-----	543
Db	1498	LLELEYEGFVYRGFFVTKKYAVIDEGKTMTRGLBIVRWDSEIAKETQARVLEALKH	1557	CC
QY	544	-----	-----	543
Db	637	DSVERAKETVKDVEELAKQVPLERKLVHFOQTDLSEKKAIGHVPAIKRLAANGKIV	696	CC
QY	544	-----	-----	543
Db	1558	GDVEAVRIVKEVTERLKVSKYEVPPKLHVHOBTRDVKYDGHVPAVAKLARGKIK	636	CC
QY	544	-----	-----	543
Db	697	RPGTISIYVLRGSKTSDRVLLSEYDSDYENQVLPAVRLTEAFGYZRKE	756	CC
QY	544	-----	-----	543
Db	1618	RPGTVIYVILGSKSRIGDRAIPFDFPAKHYDAEYYTENQVLPAVRLTEAFGYZRKE	1677	CC
QY	544	-----	-----	543
Db	757	DLKYGSSKQVGLDANLK	773	CC
QY	544	-----	-----	543
Db	1678	DLRYQTKTQVGLGANLK	1694	CC
RESULT	13			
DPOL THEORY	STANDARD;	PRT;	1668 AA.	
AC	Q9H05;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	DNA Polymerase (EC 2.7.7.7) (Contains: Endonuclease PI-PhyI (EC 3.1.1.1) (Phy pol-1 intein); Endonuclease PI-phyI (EC 3.1.1.1) (Phy pol-2 intein)) (Fragment).			
DE	(EC 3.1.1.1) (Phy pol-1 intein); Endonuclease PI-phyI (EC 3.1.1.1) (Phy pol-2 intein)) (Fragment).			
GN	Thermococcus hydrothermalis.			
OS	Thermococcales; Thermococaceae; Thermococcus.			
OC	Archaea; Euryarchaeota; Thermococcales; Thermococaceae; Thermococcus.			
OX	NCBI-TAXID=46539;			
RN	[1]			
RP	SEQUENCE FROM N. A.			
RA	Quereillou J.J.E., Cambon M.A., Leseongeur F.O., Barbier G.;			
RT	"Thermococcales taxonomy and phylogeny based on the comparative use of 16S rDNA, 16S-23S rDNA intergenic spacer and family B DNA polymerase genes.;"			
RT	Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.			
RP	[2]			
RA	CHARACTERIZATION OF PI-THYI.			
RT	MEDLINE=20512550; PubMed=11051140;			
RA	Saves I., Eleaume H., Dietrich J., Masson J.-M.;			
RT	"The Thy pol-1 intein of Thermococcus hydrothermalis is an isoschizomer of PI-THI and PI-THII endonucleases.;"			
RT	Nucleic Acids Res. 28:4391-4396 (2000).			
-1	FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE			
QY	-----			
Db	154	KGEIMIISYADEEEARVITWNIDLLYDVVSNEREMIKREVORKEKPDVYNGDN	213	CC
QY	94	PALRKTRHPRAVDITYEYDIFPAKRYLIDGLPMEGDEEKLMAFDIETFYHSCDEFG	153	CC
Db	61	PAIRDEIRRHSAVDIEYDIFPAKYLIDGLPMEGDEEKLMSFDIETLYHSGEEFG	120	CC
QY	181	EDFAYLKKRCKRIGKFTLRRGDS - EPKIORMGRFAVEVKGRHFLDLPVIRTIPL	238	CC
Db	121	TGPIMLISYADEGEARVITWKIDLPVYEVSTEKEKMRFLKVKEPDVLITYNGDN	180	CC
QY	214	FDLPLVLIKRAEKGWLLGDKRKEPEKTHRGMSFAVEKGRHFLDLPVIRTIPL	273	CC
Db	274	TYTLEAVYEAVLGKSKGAEELAIWETEESMMLQASMEDARATEYLGKEFPMEA	333	CC
Db	239	TYTLEAVYEAVFGTPKEKVEEITAWETGEGLERVARYSMEDAKVYELGREFPMEA	298	CC

QY	334 ELAKIGGQSWDVSSSSTGNLNUVWLLRAYERNELAPNKPDDEEYRRRLRTTYLGIVK 393	Db	299 OLSRLRSNLDVSSRSSTGNLNUVWLLRAYERNELAPNKPDDEEYRRRLRTTYLGIVK 393
QY	394 EPERGLWENTYLDRCLYPSITVTHNVSPTDLERBGCKNDVAPIVGKRCDFPFP 453	Db	358 EPERGLWNTVYDPMFLYPSIITHNVSPTDFNREGCKEYDTAQVGHFKDQGFP 417
QY	454 SINGELITMROEIKKKMKTADPIEKKMUDYRQRAVHLA----- 493	Db	418 SLIGCALDERQKIKKRMKASIDPLEKKLILYDQKAIIKILANSLLPBEWIPLVENGKVRHL 477
QY	494 ----- 493	Db	478 RIGERFVDKLMETDSELVLRNGDTEVLRGIRALSFRDKSKKARVYMPVKAIRHRYSGDV 537
QY	494 ----- 493	Db	494 ----- 493
QY	538 YEIVLGSGRRITVEGHSLEAYGDCELRERETGGEKAGDALLAVPRRVNLFKKERLNLVE 597	Db	598 LLRRLPEEENGDIITIPVKGRKNFKGMLRTLWNSGEEKRPRTRARRYLEHLEGIGYR 657
QY	494 ----- 493	Db	658 LKKIGYEVTDREGGLRYRKLYERLVEAARYNGNKRREYLVENAVRDVIALMPEELRDWL 717
QY	494 ----- 493	Db	718 VGTRNGFMRMPPVTEEDERKLGLGYVSEGNARKRNQKNGSYVKKLYNENQRVLDME 777
QY	494 ----- 493	Db	778 SLAERFFGRVKGKMYIEIPRKMAVIFIFENLCLGTLAENKRKVPEAIFTSPESVRWAFTEGY 837
QY	494 ----- 493	Db	838 FIGDGDVHPSKVRULSTKSELLVNLGVLWLNLSGVSAIKTRHDSVYRVVNEELPFTDV 897
QY	494 ----- 493	Db	898 RKKKNAYSHVPIPEELTEFGKPFQRSYIEKFRELVKSEKDGKAKRKEWLLNGDV 957
QY	494 ----- 493	Db	958 LDKVLEVKKRPYEGVYDLSVEEDENFLAGFGGLYAHNSYYGGYARWYCKBCEASV 1017
QY	517 TAWGRHYIEMTIKEEKGFKVLIAD----- 543	Db	1018 TAWGRDYIETIIEERFGFKVLYADSYTGETEIKRNGKVERVALERLFQRFQDVYRIG 1077
QY	544 ----- 543	Db	1078 EKEVCGVLEGVEALTLDRGLVWKSVPYWMRHTNKRKIRWFTNSWYLDVTEHDHLLIGY 1137
QY	544 ----- 543	Db	1138 MNTSKVKPGKPLKERLVEYKPGELGESVKSLITPNRAIAHGIRVNPIAVKLWELIGLVG 1197
QY	544 ----- 543	Db	1198 DGNNGGQSNWAKYVNGLSLGLDKEEIEKILKPLKNTGTSNYDKSKGDVSLSKWLA 1257
QY	544 ----- 543	Db	1258 RFMRYKIDBGSKSKRIPEMFLPREYIEFLRGLFSTADGTVSLRKGVPEVRLTSVNPTEL 1317
QY	544 ----- 543	Db	1318 SSSVRKLMLVGVNSMVFETNPNRYLGKESGTHSVHVRKDKHFAERIGFLLDRKATK 1377
QY	544 ----- 543	Db	544 ----- 543
QY	548 YATIPGERPETIKKKAKEFLKYNKNSKLPGLELEVEGFLRGFFVAKKRYAVIDEGRIT 607	Db	608 TRELLEVDRWSESAKETOKVYELKEDSVRSVERAVETYKDVWEIAYKQVPLKLVH 667
QY	668 QITKDSEYKATOPHVAIKRKLAAKGIKURPGTISIYVLRGSSKISDRVILSEYDPKK 727	Db	1498 TRGIEIVRKDWSESAKETOKAVLIRGDFVVEAVRIVKDVDTERKLISKYEVPEKLVH 1557
QY	728 HKDPDYIENQVLPAVRILEAFYGRKEDLKVOSKQVGLDAMIK 773	Db	1558 QITRELKDYKATOPHVAIKRKLAAKGIKURPGTISIYVLRGSSKISDRVILSEYDPKK 1617
QY	1618 HRVDAEYVYIENQVLPAVERILKARGIKURPGTISIYVLKGSGRIGDRAIPFEDFPTK 1663	Db	1618 HRVDAEYVYIENQVLPAVERILKARGIKURPGTISIYVLKGSGRIGDRAIPFEDFPTK 1663
RESULT 14			
DPOL	METVO	DPOL	METVO
ID	DPOL METVO	STANDARD	PRT;
AC	P52025;		824 AA.
DT	01-OCT-1996 (Rel. 34, Created)		
DT	15-DEC-1998 (Rel. 37, Last sequence update)		
DE	DNA polymerase (EC 2.7.7.7).		
GN	POL.		
OS	Methanococcus voltae.		
OC	Archaea; Eurarchaeota; Methanococcales; Methanococcaceae;		
OC	Methanococcus.		
OX	NCBI_TAXID:2188;		
RN	{1}		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE-95014087; PubMed-7929013;		
RA	Konisky J., Paule S.M., Carnato M.E., Kansy J.W.;		
RT	"the DNA polymerase gene from the methanogenic archaeon Methanococcus voltae.",		
RL	J. Bacteriol. 176:6402-6403(1994).		
CC	-!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate + (DNA)(N).		
CC	-!- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.1sb-sib.ch/announce/ or send an email to license@1sb.sib.ch).		
CC	-----		
DR	EMBL: L33366; AAAT2443_1; -.		
DR	HSSP; P56689; ITGO; InterPro; IPR020204; DNA_pol_B.		
DR	Pfam: PF00136; DNA_pol_B; 1.		
DR	Pfam; PRO3104; DNA_pol_B_exo; 2.		
DR	PRINTS; PRO00106; DNAPOLB.		
DR	SMART; SM00486; P0LBC; 1.		
DR	PROSITE; PS00116; DNA_polymerase_B; 1.		
KW	PROSITE; PS00116; DNA_polymerase_B; 1. DNA-binding; DNA-directed DNA polymerase; DNA replication; DNA-sequence.		
SQ	SEQUENCE 824 AA: 96754 MW: 094A630D8C1514FC CRC64:		
QY	Query Match 34.6%; Score 1392; DB 1; Length 824; Best Local Similarity 36.1%; Pred. No. 4.2e71; Mismatches 308; Conservative 179; Mismatches 251; Indels 116; Gaps 21;		
QY	4 DTPIYTKGKPTIRIFKENG---EFTKELDRHFQPTYALKDSDA-----D 49		
QY	2 DLDYNSKD-LCIDMYKNCKGLKPEINLQKECEFKPFYVDTSEPEKTYDLDGLNQEI 59		
QY	-----TDGF 547		

Qy 393 KEPERGLWENTYLDFRCLYPSITVHNSPDTLREGCKN-YVAPIVGKFCDFPG 450
 Db 397 LEPERGLHENYACLDPASMTPSIMITNSPDTY--GCRRDCYE-APEVGHKERSPDG 452
 Qy 451 FIPSILGELTMWQETKKMK-ATIDPIEKMMDFQRAYKLHANSYGGMYPKARWS 509
 Db 453 FFKRILMLJEKREKVELKLNLSSESSEYKLDITKQQTUKVLNTNSFYGMGNLARWC 512
 Qy 510 KECABESVTANGRHYTEMTEKEEKGFKLYADIDGFYATIPGKPKETKKAKEFLK 569
 Db 513 HPCABATTANGRRHFTTSK-TAESMGFKVLYGDTDSIFVTKAGMPTKEDVDR---LIDK 567
 Qy 570 INSKUGLLELNEFYLROFFVAKRYAVIDEGRITRGLLEVRRDSEIATEQAV 629
 Db 568 LHEELP--TQIEVDEVYSAIFFVAKRYAGLTEDGRLVKGLEVRGDMCELAKKVOREV 625
 Qy 630 LEAIKEDSVEKAVETVKDVVEIAKYQVPLEKLYTHEQITKDUSEYKATGPVHALAKRL 689
 Db 626 IEVILKEKNBKAISLKVDTLRIKEGVSKLEVVIVKGTKESKYESMQAHVAALKA 685
 Qy 690 AAKGIRVRPGTISTIVLRSKGKTSDR--VILUSEYDPKK-----HKYDPDYY 735
 Db 686 REMGIIYPVSSKIGVIVKGSGNIGDRAYPIDLIEDFDGENLRKTKSGIEKKLKDYY 745
 Qy 736 IENQVLPAVLILEAFGYREDLKQSSKGQVGLDAW 771
 Db 746 IDNQIIPSPVRLERFGYTESLK--GSSQMSLDSF 779

Search completed: May 28, 2002, 14:22:19
 Job time: 252 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model.

Run on: May 28, 2002, 14:17:02 ; Search time 23.89 Seconds

Sequence: (without alignments)
1 MIFDTDYITKDGKPIIRFK.....KEDLKYQSSKQVQGIDAWLKK 774

Title: US-09-803-165-34
Perfect score: 4026

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Scored: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database : PIR 71:*

1: pirl:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Query Match length DB ID

Description

RESULT 1

S42459 DNA-directed DNA polymerase (EC 2.7.7.7) vent, intein containing precursor - Thermoco

N; Contains: DNA endonuclease (EC 3.1.-) PI-I; DNA endonuclease (EC 3.1.-) PI-II;

C; Species: Thermococcus litoralis

C; Date: 31-Dec-1993 #sequence revision 02-Aug-1994 #text_change 18-Jun-1999

C; Accession: S42459; S42451; S42450; S42458

R; Perler, F.B.; Comb, D.G.; Jack, W.E.; Moran, L.S.; Oiang, B.; Kucera, R.B.; Benner,

submitted to the EMBL data library, September 1992.

A; Reference number: S42459

A; Cross-References: EMBL:MT198; NID:9154685; PID:AAK7200.1; PID:9154686

R; Perler, F.B.; Comb, D.G.; Jack, W.E.; Moran, L.S.; Oiang, B.; Kucera, R.B.; Benner,

Proc. Natl. Acad. Sci. U.S.A. 89, 5577-5581, 1992

A; Molecular type--DNA

A; Residues: 1-1702 <PERI>

A; Accession: S42459

A; Reference number: S42458

A; Cross-References: EMBL:TM74398

R; Hodges, R.A.; Perler, F.B.; Noron, C.J.; Jack, W.E.

Nucleic Acids Res. 20, 1613-6157, 1992

A; Title: Protein splicing removes intervening sequences in an archaea DNA polymerase.

A; Reference number: S40788; MURID:93117083

A; Contents: annotation

C; Function: <END2>

A; Description: nucleotidyltransferase

A; Note: DNA-directed DNA polymerase Vent

C; Function: <END1>

A; Description: endonuclease

A; Note: DNA endonuclease PI-III

C; Function: <END2>

A; Description: endonuclease; hydrolase

A; Note: DNA endonuclease PI-III

C; Function: <END1>

A; Description: endonuclease; hydrolase

A; Note: DNA endonuclease PI-III

C; Function: <END2>

A; Description: endonuclease; hydrolase

A; Note: DNA endonuclease PI-III

C; Superfamily: Thermococcus litoralis DNA-directed DNA polymerase Vent

C; Keywords: DNA binding; endonuclease; hydrolase; nucleotidyltransferase; protein spl

F; 1-494; Domain: DNA-directed DNA polymerase Vent extein 1 #status predicted <XT1>

F; 495-1032; Product: DNA endonuclease PI-III (pol vent intein 1) #status predicted <XT2>

F; 1033-1081; Domain: DNA-directed DNA polymerase Vent extein 2 #status predicted <XT2>

F; 1082-1471; Product: DNA endonuclease PI-I (pol vent intein 2) #status predicted <XT2>

F; 1472-1070; Domain: DNA-directed DNA polymerase Vent extein 3 #status predicted <XT3>

F; 494-1033; Cross-link: peptide (Asn-Ser) #status predicted <XT3>

F; 1081-1472; Cross-link: peptide (Asp-Thr) #status predicted

ALIGNMENTS

30 605.5 15.0 1106 1 A39299 DNA-directed DNA p
31 587 14.6 1492 2 T18560 DNA-directed DNA p
32 584.5 14.5 879 1 A5677 DNA-directed DNA p
33 583.5 14.5 1462 1 DJHUAC DNA-directed DNA p
34 581 14.4 1081 2 T20698 DNA-directed DNA p
35 578.5 14.4 875 2 JC5186 DNA-directed DNA p
36 570.5 14.2 959 2 F72763 DNA-directed DNA p
37 570 14.2 1465 2 S45628 DNA-dependent DNA p
38 559.5 13.9 223 2 E69125 DNA-directed DNA p
39 548.5 13.6 1339 1 S20052 DNA-directed DNA p
40 527.5 13.1 1468 2 S58220 DNA-directed DNA p
41 522.5 13.0 1505 2 S28079 DNA-directed DNA p
42 516.5 12.8 844 2 T31321 DNA-directed DNA p
43 516 12.8 1405 1 DJZPA DNA-directed DNA p
44 504.5 12.5 2154 2 A84669 hypothetical prote DNA-directed DNA p
45 500.5 12.4 913 2 T17675 DNA-directed DNA p

QY	1	MIFDTDTYTKDGSPIRTRKKENGKFEKIDLPHQPIYALLKDDSADEIKAKRGEHG 60	Db	1081	DSVSESEETIQRQNKIREVKIKDLSKVDISIGECYCLEVERALTLDDGKVWKPV 1140
Db	1	MILDITYTKDGPPIRTRKKENGKFEKIDLPHQPIYALLKDDSADEIKAKRGEHG 60	Qy	544	-----
Qy	61	KIVRVDAVKVKFLGFDPEWVKLIFBHQDPAKGKIREPAVIDEYDIPFARRY 120	Db	1141	PYMRHRANKRMFRWLNTNSWYIDVTDHSLSLGYINTSKTKTAKRIGERLKEVKEFELGR 1200
Db	61	KIVRVDAVKVKFLGFDPEWVKLIFBHQDPAKGKIREPAVIDEYDIPFARRY 120	Qy	544	-----
Qy	121	LIDKGLPMEGDEELKJMAFDIETFYHGGDFGKGEIMISVADEEARVITWNKIDLY 180	Db	1201	AVKSLICPNAPLKDNTKTSEIAVFKWELVGLIVDGNNNGGDSRWAEEYVGLSTGKDAE 1260
Db	121	LIDKGLPMEGDEELKJMAFDIETFYHGGDFGKGEIMISVADEEARVITWNKIDLY 180	Qy	544	-----
Qy	181	VDWNSNREMIKRFQVTRKEPPDVLTYNGNFDPLYLIKRAEKGTVLGLGRDKEPE 240	Db	1261	IKRKLLEPLKTYGVISNYPPNEKDFNLAKSLVKFMKRFKDEGRKRIPEFWELPV 1320
Db	181	VDWNSNREMIKRFQVTRKEPPDVLTYNGNFDPLYLIKRAEKGTVLGLGRDKEPE 240	Qy	544	-----
Qy	241	PKTRKGDSFAVEIKGRHFDEPWPVRTINLPTYLEAVVAVLGKTKSKLGAEIAI 300	Db	1321	TYIEFLRLSADCTVTIRKGVPBIRLNTIDADFLREVRKLWIVGINSIAETTPNR 1380
Db	241	PKTRKGDSFAVEIKGRHFDEPWPVRTINLPTYLEAVVAVLGKTKSKLGAEIAI 300	Qy	544	-----
Qy	301	WETEESMKKLAQYSMEDARAT'ELGKEPPMEEALAKLIGQSYWDVSRSTGNLVEWLL 360	Db	1381	YNGVSTGTVSKHLRKKNKRAERIGFLIERKQKRLLEHKSARWKRNIDFGFLDVHK 1440
Db	301	WETEESMKKLAQYSMEDARAT'ELGKEPPMEEALAKLIGQSYWDVSRSTGNLVEWLL 360	Qy	544	-----
Qy	361	RVAYERVELLAPKPDDEBRYRRLRTTYLGYYKEPERGLWENITYLFRCLPSITVHN 420	Db	1441	KVEETPYEGWVDEVEEHTFRANILWHTDGTATIFGERKPELKKAKERLAIN 1500
Db	361	RVAYERVELLAPKPDDEBRYRRLRTTYLGYYKEPERGLWENITYLFRCLPSITVHN 420	Qy	573	KLPDLLELVEYEGFLRGFVAKKAVDIEBGRITRGLFVWRDWESETAKTOVKLEA 632
Qy	421	VSPDTLREGCKNYDAPIVGKFCPPGFPSPSILGELITMRQEIKKKMKTIDPEKK 480	Db	1501	KLPGILEVEYEGFLRGFVFTKRYAVIDEGRITRGLFVWRDWESETAKTOVKLEA 1560
Db	421	VSPDTLREGCKNYDAPIVGKFCPPGFPSPSILGELITMRQEIKKKMKTIDPEKK 480	Qy	633	ILKEDSVRAVEVTKVDELAQKQVPLKLVHQTDLSEKAIGHVAIKRLAK 692
Qy	481	MUDYRQRAKLLANSILPNEWIPPIENGKEIKFVKIGEFINSTMEMQENKVVENTEVL 493	Db	1561	ILKGDSVSEKAVEVTRDWEVKIYRPLKLVHTHQITRDLKQKIGHVATKRLAAR 1620
Db	481	MUDYRQRAKLLANSILPNEWIPPIENGKEIKFVKIGEFINSTMEMQENKVVENTEVL 493	Qy	693	GKVRPGTISIYTIVLGSKISDRVILSSEYDPKPKHVKYPPDYIENQVLPAVLITLEAFG 752
Qy	494	-----	Db	1621	GKVRPGTISIYTIVLGSKISDRVILSSEYDPKPKHVKYPPDYIENQVLPAVLITLEAFG 1680
Db	494	-----	Qy	753	YRKEDLYQSSKQGDAMLUK 774
Db	494	-----	Db	1681	YRKEDLYQSSKQGDAMLUK 1702
Db	601	KEVSGDGKEGLIVAPKKIKLNEGVYSINIPELISDSEETADYMTISAKGRKRNFF 660	RESULT	2	-----
Qy	494	-----		493	-----
Db	661	GMLRTLRRMFGEEENRRIRTBNYLFHLEKLGLLIKLLPRGYEVTDWERLKYKOLYK 720		67920	-----
Qy	494	-----		493	DNA-directed DNA polymerase (EC 2.7.7.7) - Thermococcus sp.
C	721	SVKYNGNKRKEYLVMFNETKDFISYFPOKELEEMKIGTLNGFRNTCLKVDEDFGKLLGY 780	C;Species:	Thermococcus sp.	
C	494	-----	C;Date:	17-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 18-Jun-1999	
C	494	-----	C;Accession:	S67920	
C	494	-----	R;Southworth, M.S.; Kong, H.; Kucera, R.B.; Ware, J.; Jannasch, H.W.; Perler, F.		
C	494	-----	submitted to the EMBL data Library, January 1996		
C	781	VSEGYAGAQKNTGGISYSVLYNEDPNVLESMKVNAEKEFGKVRVDRNCVSKMAYL 840	C;Description:	Cloning, expression, and modulation of the 3'-5' exonuclease activity	
C	494	-----	A;Reference:	567920	
C	494	-----	A;Accession:	S67920	
C	494	-----	A;Molecule type:	DNA	
C	494	-----	A;Residues:	1-775 <SOU>	
C	494	-----	A;Cross-references:	GB-U47108; NID:91197451; PID:AAA88769.1; PID:91197452	
C	841	VMKCLCGALAAENKRIPSVILISPEPVWMSFLEAYFTGDDIHPSKRFLSTKSELLANOL 900	C;Experimental source:	strain DNA-directed DNA Polymerase	
C	494	-----	C;Superfamily:	herpesvirus DNA-directed DNA Polymerase	
C	901	VFLNSLGISSSVKIGFDSGVYVYINEDLQFOTSREKNTYSLIPKIELRUVFGKEFQ 960	C;Keywords:	DNA binding; nucleotidyltransferase	
Qy	494	-----	Query Match	79.6%	Score 3204.5; DB 2; Length 775;
Db	961	KNMTPKKFKELVDSGKUNREKAKLEFFINGDILVDRVKSVKEKDVGCGVYVOLSENEEN 1020	Best Local Similarity	76.6%	Best Local Similarity 76.6%; Pred. No. 2.5e-157;
Db	961	KNMTPKKFKELVDSGKUNREKAKLEFFINGDILVDRVKSVKEKDVGCGVYVOLSENEEN 1020	Matches	592;	Matches 592; Conservative 89; Mismatches 89; Indels 3; Gaps 2;
Qy	1	MIFDTDTYTKDGSPIRTRKKENGKFEKIDLPHQPIYALLKDDSADEIKAKRGEHG 60	Qy	1	MIFDTDTYTKDGSPIRTRKKENGKFEKIDLPHQPIYALLKDDSADEIKAKRGEHG 60
Qy	494	-----	Best Local Similarity	76.6%	Best Local Similarity 76.6%; Pred. No. 2.5e-157;
Qy	494	-----	Matches	592;	Matches 592; Conservative 89; Mismatches 89; Indels 3; Gaps 2;
Db	1	MILDITYTKDGPPIRTRKKENGKFEKIDLPHQPIYALLKDDSADEIKAKRGEHG 60	Db	1	MILDITYTKDGPPIRTRKKENGKFEKIDLPHQPIYALLKDDSADEIKAKRGEHG 60
Db	1021	FLVFGFLGLYAHNSYVGYNGYKPARWSKECAESVTANGRHYIEMTIREBIEKEFGKVL 1080	Qy	61	KIVRVDAVKVKFLGFDPEWVKLIFBHQDPAKGKIREPAVIDEYDIPFARRY 120
Qy	543	D-----	Qy	543	-----
Qy	543	-----	Qy	543	-----

F:1-775/Product: DNA-directed DNA polymerase #status experimental <MAT>						
Query Match	78.3%	Score	3151	DB	2	Length
Best Local Similarity	74.6%	Pred. No.	1.4e-154			
Matches	576;	Conservative	100;	Mismatches	94;	Indels
C:Species	Pyrococcus sp.					
C:Accession	S68593					
C:Description	In vitro protein splicing of purified precursor and the identification of					
A:Accession	S68593					
A:Molecule type	DNA					
RESULT	5					
668593						
DNA-directed DNA polymerase (EC 2.7.7.7) vent, intein containing precursor - Pyrococcus						
N:Contains: DNA endonuclease (EC 3.1.-.-) PI-1; DNA-directed DNA polymerase (EC 2.7.7.7)						
C:Submits to the EMBL Data Library, August 1993						
C:Date: 24-Aug-1996 #sequence-revision 01-Nov-1996 #text_change 24-Sep-1999						
C:Accession: S68593						
R: Xu, M.Q.; Southworth, M.W.; Marsh, F.B.; Hornstra, L.J.; Perler, F.B.						
R: Description: In vitro protein splicing of purified precursor and the identification of						
A:Accession: S68593						
A:Molecule type: DNA						

A; Residues: 1-1312 <XUR>
A; Cross-references: EMBL:000707; NID:9436192; PIDN:AAA61130_1; PID:9825735
R; Xu, M.Q.; Southworth, M.W.; Marsha, F.B.; Hornstra, L.J.; Perier, F.B.
Cell 175, 1371-1377, 1993
A; Title: In vitro protein splicing of purified precursor and the identification of a
A; Reference number: S88581; MUID:94094330
A; Contents: annotation
A; Function: <NUCL>
A; Description: nucleotidyltransferase
A; Note: DNA-directed DNA polymerase vent
C; Function: <ENDO>
A; Description: endonuclease; hydrolase
C; Superfamily: hypothetical protein PH0202
C; Keywords: DNA replication; endonuclease; hydrolase; nucleotidyltransferase; protein
F1-1-491_1030-1312; Product: DNA-directed DNA polymerase Vent #status predicted <XT1>
F1-1-411_Domain: DNA-directed DNA polymerase Vent extein 1 #status predicted <XT1>
F1-92-1029_Product: DNA endonuclease P1 (pol I) (pol I) (pol I) #status predicted <XT1>
F1-92-1312_Domain: DNA-directed DNA polymerase Vent extein 2 #status predicted <XT2>
F1-491-1030/Cross-link: peptide (Ala-Ser) #status experimental

Query Match 72.9%; Score 2934.5; DB 2; Length 1312;
Best Local Similarity 45.6%; Pred. No. 3.8e-13; Matches 597; Conservative 86; Mismatches 87; Indels 539; Gaps 3;

Qy. 1 MIEDDTMTKDGKPLIRFEKKENGEFKIEDDHQFQWYIALKDDSAIDEKAICERHS 60
Db 1 MIDDADYIETEDGKPLIRFKKENGEFKYEDNRFPYIYALKDDSQIDEVKITAERG 60

Qy 61 KTVRVVDAVKVKKFLGRDVEVWLKTFEPHQDPVPAKKGKTRBPAVIDYEDIPRAKY 120
Db 61 KIVRILDEKVRKKFLGRPIEWVWLTFEPHQDPVPAKDKRKREHSAYVIDEYDIPRAKY 120

Qy 121 LIDKGLIPMEGDEELKIMAFDTEFYIEGDEFGKGELIMISADEEEARVITWKNDLPP 180
Db 121 LIDKGLIPMEGDEELKLLAFDITYLHGEEFAKGPIIMISADEEBAKIVTWKKDLPY 180

Qy 181 VDVSNEREMIRKFRVQVLUKEPDKVWDLVYDQFPLYLIRRAKGVTLGRKHEP 240
Db 181 VEVYSSEREMIRKFLKVIREKOPDVITYNGDSFDQDYLVRKRAKGKPLGLRGS--E 238

Qy 241 PKTHRMGDPSFAVEITKGRTHDFLDLPPVVRTINPTTYLEAVVTEAVLQTKSKGAETIA 300
Db 239 PKMQRGIDMTAIEKGRHFDYKHWIRTINPFTYLEAVVTEAVLQTKPKVYHIAA 298

Qy 301 WEEIEESMKKLAQYSMEDRARATYELGKRFPPMAELAKLIGSYWYDVSRSRSTGNLYEWYL 360
Db 299 WEGKGKLERVKYSMEDAKVYDVTYBLGRERFPMEQALQSLVQGQLWDYRSRSTGNLYEWYL 358

Qy 351 RVAYERNLIAPKNPKDDEEYRRLLTIGSYKEPERGLMITYDFRCRCPYLSIVTHN 420
Db 359 RKAYERNEALAPNPKDDEEYRRLLTIGSYKEPERGLMITYDFRCRCPYLSIVTHN 418

Qy 421 VSPDTLREBGCKNDVADIVGKFKCOPGPFTSITGELITMROEIKKKMATAIDIEK 480
Db 419 VSPDTLREBGCKNDVADIVGKFKCOPGPFTSILKRLDQEIKRKMAKSPDIEK 478

Qy 481 MLDYQRQWV----- 490
Db 479 MLDYQRQAKILANSLPPEWYPLIKNGKVFRIGDFVGLMKANQKVKTGDTVE 538

Qy 491 ---LHA----- 493
Db 539 VNGTHAFSDRKKARVMKAVTRHRYSGNVYRIVLNSGRKITITEGHSLFVYRNGD 598

Qy 494 ----- 493

Db	599	VEATGEDVKIGDLAVPRSVNLUPEKKERLNLIVELLNLISPEETEDILITIPWGRKNNFK	658
Qy	1	MFDPPYTRDGKPIRIFRKENGSEKIELPHFOPPYIAALKDOSAIDPRAKAIKGERHG	60
Qy	494	-----	-----
Db	1	-----	493
Db	659	GMLRTLRWIFGEEKVRVTASRYLRHLENLGYTRLRKIGYDIDKEGLKRYTLYEKLV	718
Qy	494	-----	-----
Db	719	VRYNGNKRREYLVEFNAVRDVISLMPPEELKEWRIGTRNGFRMGTEVFDIDEDFAKLIGYV	778
Qy	494	-----	-----
Db	779	SEGSARKWKKNQTTGGWSITVRLYNENDEVLDDMEHLAKKFFGKVKGKRNIVEPKKMAYI	838
Qy	494	-----	-----
Db	839	FESLCGTLAENKRVPPEVFTTSKGVRWAFLLEGFIGDGDVHPSKVRVLSTKSBLVNGLV	898
Qy	494	-----	-----
Db	949	-----	493
Db	899	LLNSLGSVSAIKLGVDGSGVYRVVNEELKFTYRKKNVYHSHIVPKDILKEIFGKVFO	958
Qy	494	-----	493
Db	959	NIYKKFRELVENGKLDREKAKRIEWWLGDVLDVKEVRLYDGVYDVSDEDEFN	1018
Qy	494	-----	-----
Db	1019	LAGFGFLYAHNSYGGYGAKRWCYKECAESVTAWGRYIEFVRKELEKFGFKVLYID	1078
Qy	544	TDFGYATIPGEKPETIKKKAKFELKYKNSKLPGLLEYEYEGYLRGFVAKKRYAVIDE	603
Db	1079	TDGLYATIPGARPEEIKKKALEFVDYNAKLPGLEEYEGTVRGFVTKKYLIDEE	1138
Qy	604	GRTITRPLEVVRDWESEIAKETOQAKVLEAIIKEDSVKAVELVTDVETIAKQVPLE	663
Db	11139	GKIKTRGLEIYRDSSEIAKETOQAKVLEAIIKRGHNVFAVKVKEFEKSKXIEPPKL	1198
Qy	664	VIHQITKDLSEYKAATGPHVATKRLAAGTKVPRGTLISYVLRGSGKISPRVILSEY	723
Db	11199	VIYEQTTRPLRHEYKAATGPHVAKRLAARGVVKRPGMVIGYIVLRLQDGPISKRAILABEF	1258
Qy	724	DPKKHKDPPDYYENQVLPAVRILEAFGRYKEDLKIQSSKQVGLDawl	772
Db	1259	DLRKHKYDAEYVYIENQVLPAVRLIEAFGRYKEDLRLQKTKOTGLTAWL	1307
RESULT	6		
C71210	probable DNA-directed DNA Polymerase - Pyrococcus horikoshii		
C;Species	Pyrococcus horikoshii		
C;Date	14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000		
C;Accession	C71210		
R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekine, M.; Ohfuki, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kusilida, N.; Oguchi, DNA Res. 5, 55-56, 1998			
A;Title	Complete sequence and gene organization of the genome of a hyper-thermophilic a		
A;Accession	C71210		
A;Status	preliminary; nucleic acid sequence not shown; translation not shown		
A;Molecule type	DNA		
C;Cross-references	GB:AP00007; NID:93236134; PID:BAA31074.1; PID:93258391		
A;Experimental source	strain OT3		
A;Note	this accession replaces an interim accession for a sequence replaced by GenBank		
C;Genetics			
A;Gene	PH947		
C;Superfamily	hypothetical protein PH0202		
Query Match	71.7%	Score 2886; DB 2; Length 1235;	
Best Local Similarity	46.5%	Pred. NO. 1.1e-140	
Matches	574;	Conservative 107; Mismatches 89; Indels 464; Gaps 5;	

QY	620	EIAKEOAKVLEIAKEDSEKAVETVKDVEEIAKYOPIEKLKIHETOAKDSEYKAI	679			
Db	1078	EIAKEOAKVLEIAKEDSEKAVETVKDVEEIAKYOPIEKLKIHETOAKDSEYKAI	679			
QY	680	GPHVATKRAAKGKIVRPTISIVLRSRKISDRVILSEYDPRKKHYPDVIENQ	739			
Db	1138	GPHVATKRAAKGKIVRPTISIVLRSRKISDRVILSEYDPRKKHYPDVIENQ	739			
QY	740	VLPAVIRILEAFGYRKEDLKVQSSKQVGLDAWLK	773			
Db	1198	VLPAVIRILKAFGYKREDLRNQTKVQVGLGAWIK	1231			
RESULT	7					
	571551					
Db		DNA-directed DNA polymerase (EC 2.7.7.7.) KOD, intein containing precursor - Pyrococcus				
QY		S				
Db		Contains: DNA endonuclease (EC 3.1.-.) PI-I; DNA endonuclease (EC 3.1.-.) PI-II; DNA				
C; Species:		Pyrococcus sp.				
A; Variety:		strain KOD1				
C; Accession:		S71551				
C; Date:		22-Nov-1995 #sequence_revision 22-Nov-1995 #text_change 20-Apr-2000				
C; Kakihara, H.; Takagi, M.; Inanaka, T.		submitted to the EMBL Data Library, March 1994				
A; Description:		Cloning and analysis of the DNA polymerase gene from a new hyperthermophilic				
A; Reference number:		S71551				
A; Accession:		S71551				
A; Molecule type:		DNA				
A; Residues:		1-670 <ENAK>				
A; Cross-references:		EMBL:D29671				
C; Function:		<NNCL>				
A; Description:		as DNA-directed DNA polymerase KOD				
C; Note:		DNA-directed DNA polymerase KOD				
C; Function:		<EN1>				
A; Description:		as DNA endonuclease PI-I, catalyzes the hydrolysis of internal phosphodi				
C; Note:		as DNA endonuclease PI-PsP1				
C; Function:		<EN2>				
A; Description:		as DNA endonuclease PI-II, catalyzes the hydrolysis of internal phosphodi				
C; Note:		as DNA endonuclease PI-PII				
C; Superfamily:		DNA-directed DNA polymerase KOD				
C; Keywords:		endonuclease; hydrolase; nucleotidyltransferase; protein splicing				
F; 1-406, 679-851, 1384-1670; Product:		DNA-directed DNA polymerase KOD #status predicted <MPR>				
F; 1-106; Domain:		DNA-directed DNA polymerase KOD extein 1 #status predicted <XT1>				
F; 407-766; Product:		DNA endonuclease PI-I (pol KOD intein 1) #status predicted <XT2>				
F; 857-851; Domain:		DNA-directed DNA polymerase KOD extein 2 #status predicted <XT3>				
F; 1388-1670; Product:		DNA endonuclease PI-II (pol KOD intein 2) #status predicted <XT4>				
F; 406-767; Cross-link:		DNA-directed DNA polymerase KOD extein 3 #status predicted <XT5>				
F; 851-1384; Cross-link:		peptide (Arg-Ser) #status predicted <XT6>				
Query	Match	68.7%; Score 2764.5; DB 2; Length 1670;				
Best Local Similarity		35.9%; Pred. No. 2.9e-334;				
Matches	599;	Conservative	83;			
	Mismatches	88;	Indels	899;	Gaps	4;
QY	1	MIFPDYDITKOCKPPIRIFKENGKENGKIEDLPHPFOPIYIAALKDSAIDEKAINGERHG	60			
Db	1	MILDPTDYDITKOCKPPIRIFKENGKENGKIEDLPHPFOPIYIAALKDSAIDEKAINGERHG	60			
QY	61	KIVRVDVDAVKKKFGLGRDVEWKLIFEEHPPDOPALRKREHPAVIDYDIFAKRY	120			
Db	61	IVTIVKVRKEVKKELGRPEVWKLKFTHQDYPAIRDKREHGAVIDYDIFAKRY	120			
QY	121	LIDKGGLIPMDEEKLMAPIDETYHHEGFBGKGEIMMSYADEEARYTWTKNIDPV	180			
Db	121	LIDKGGLIPMDEEKLMAPIDETYHHEGFBGKGEIMMSYADEEARYTWTKNIDPV	180			
QY	181	WDVWSNEREMIKRFVIVRKEPDVITYNGDNFDIPLYLKRAEKIGVTLIGRKEHPE	240			
Db	181	WDVWSNEREMIKRFVIVRKEPDVITYNGDNFDIPLYLKRAEKIGVTLIGRKEHPE	240			
QY	241	PKIHRMGDSFAVEIKRIFHDLPFYRVRTINLPTYLEAVIGTKTKRGLGAEIAAI	300			
Db	239	PKIQRMGDRFAVEVKRIFHDLPFYRVRTINLPTYLEAVIGQPKVKYABITPPA	298			
QY	301	WETEESKKLQDQYSMERAQARYELGEFFPMERELAKLIGOSWWDYRSRSGNQVWELL	360			
Db	299	WETGENIERYARYSMDAKQYELGEFLPHEAQSLRSLIGOSLWDVSRSQGNTVWELL	358			
QY	361	RVAYERNELAPNPKPDEBEYRRLRTTYLGGYVKERPERGLWNTYDFRC	410			
Db	359	RKAYERNELAPNPKPDEKELARP-ROSIEGGIVKEPERGLWNTYDFRC	417			
QY	411	-	-			
Db	418	GKGININSEVOEGDYYLGIDGWQRVRKVWEDYKGEVLYNINGLKCTPNHKLPVVTNERQ	477			
QY	411	-	-			
Db	478	TRIRDSSLAKSFLTKVKGKIIITPLYEIGRATESNIPPEEVVLKGELAGILLAESTLRLK	537			
QY	411	-	-			
Db	538	DVEFDSRKRRKISHQYRVEITGDEEERDRITIYFERLEGIFPSISRKGMIAVL	597			
QY	411	-	-			
Db	598	KVAKKNVLYKEIMDNIESLHAPSVLRGFFBGGDSVNVRVRSIVATQTKNEWKIKLVS	657			
QY	411	-	-			
Db	658	KLSQLQGIPHQTYQYQENGKDRSRVLEITGKGDLQTLIGFISERKNALLKAI	717			
QY	411	-	-			
Db	718	QREMNLENNGFYRLSEPVNSTEYVKGKVIDITLETPYFANGIITHNSLXPSITH	777			
QY	421	VSPDTLREGKKNYDAPIVYKFCDFPGFIPSIGELITMROEKKKMATIDIEKK	480			
Db	778	VSPDTLNREGCKEYDVAPOVGHFRCKDFPGFIPSLLDLEBROKIKKKMATIDIEKK	837			
QY	481	MIDYRQRAIKLANSIPEENLVPVLEGEVHVRIGELIDRMMEENAGVKREGETEVLE	493			
Db	838	LIDYRQRAIKLANSIPEENLVPVLEGEVHVRIGELIDRMMEENAGVKREGETEVLE	897			
QY	494	-	-			
Db	898	VSGLEWPSFNRTNKAELKRVKALIRDISGVWTIRLKSRRIKTSGHSLSFVNGEL	957			
QY	494	-	-			
Db	958	VEVTGDELKPGDLYAVPRRLELPERNHVLNLVELLIGTPEEBETDWTIPVKGKNNFK	1017			
QY	494	-	-			
Db	1018	GMILRTLWIFGEKRPRTARRYRLHEDLGYVRLKIGYEVLDWDSILKNYRRLYEALVEN	1077			
QY	494	-	-			
Db	1078	WRYNGKREYLVEPNSIRDAGIMPKEKIGLNGFRMKLIEVDESLAKLGYV	1137			
QY	494	-	-			
Db	1138	SEGYARKORNPKNGWSVSVKLYNEDPEVLDIMERLASREFGKVRRGRNRYVIEPKKIGYL	1197			
QY	494	-	-			
Db	1258	LINSVGSVSAVKLGHDSGVYVYINELLPFVKLKKKNAVSHVPIEVLSFVGKFQKN	1317			
QY	494	-	-			
Db	1198	FENMCGVLAENKRIPPEVFTSPKGVLAFLEGYSSAMTSQEIQALNEKRALANQLVL	1257			
QY	494	-	-			
Db	1318	VSPQTTRKMDVEDGRUPPEKAORLSSWILLEGDVVLDRVRESVDVVEDYDQYVYDLSVEDNNFL	1377			

QY 494 -----NSYGYGMGPKARYSKECESVWAWGRHYTEMTEKEERFGKVLYADT 514
 ||||| ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 1378 VGRGLIVYAHNSYYGGYARWYCKECAESVWAWGRHYTEMTEKEERFGKVLYADT 1437
 QY 545 DGFATIGADAEVKKKAMEFLNYINAKLPGLELEFGYRGGFFTKKKAVIDEQ 604
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 1438 DGFATIGADAEVKKKAMEFLNYINAKLPGLELEFGYRGGFFTKKKAVIDEQ 1497
 RIM"RGLVERRDSWEIKAETQAVLEATKESVAKVLEAFYRGGFFTKKKAVIDEQ 664
 QY 605 RIM"RGLVERRDSWEIKAETQAVLEATKESVAKVLEAFYRGGFFTKKKAVIDEQ 664
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 1498 KITRGLLEIVRDRSEIKAETQAVLEAFYRGGFFTKKKAVIDEQ 1557
 QY 665 IHQQTOKDSEYKAIGPHVIAKRLAAGKIGVPGTISYIVLRSKGSKISDRVLLSYD 724
 ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 1558 IHQQTDLKDYKATGPVKRVLAAVGKRPGVTSIVLKGSGRIGDRAPEDEQ 1617
 QY 725 PKHKKDDYYIENQVLPVLRLAFGKRYKEQDLYQTKRQVLSAWLK 773
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 b 1618 PTKHKYDAEYYIENQVLPVLRLAFGKRYKEQDLYQTKRQVLSAWLK 1666
 QY 726 PTKHKYDAEYYIENQVLPVLRLAFGKRYKEQDLYQTKRQVLSAWLK 1666
 RESULT 8
 A69312 DNA polymerase B1 (polB) homolog - Archaeoglobus fulgidus
 C;Species: Archaeoglobus fulgidus
 C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C;Accession: A69312
 R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.E.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.; Nature 390, 364-370, 1997
 A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.; Smith, H.O.; Woese, C.R.; Venter, J.C.
 A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon A;Accession: A69312
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-781 <KLE>
 A;Cross-references: GB:AE001070; GB:AE000782; NID:92689393; PIDN:AAB90741.1; PID:9265013; C;Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match 32.3%; Score 1301; DB 2; Length 781;
 Best Local Similarity 37.4%; Pred. No. 1.9e-59;
 Matches 305; Conservative 155; Mismatches 266; Indels 90; Gaps 21;

Y 2 IFFDTDITKDGKPIRFFKKEENQEFKIELDPHDQPYIYALLKDSAIDEIKATKGERHCK 61
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 8 LIDDADEPIGKAVWRLACKDGGIFVAYDYNQDPYFVIGDDE--DILKNATSTR 64
 QY 62 IYRVVDAVKKFKLGRDVEWVNLIFEPHQRDVALRKIREHAPVIDEYDIPFAKYL 121
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 65 VIKLKSFEKAQLKTLGREVEGYIYAHFOHFKLRLYLSQFG --DVERADIPFAKYL 121
 A;Start codon: ATG
 C;Function: <DP>
 A;Map Position: REW816304-811400
 A;Genetics:
 A;Cross-references: GB:U67532; GB:L77117; NID:91591559; PID:91591563; TIGR:MG0885; PI
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-1634 <BUL>
 A;Cross-references: GB:U67532; GB:L77117; NID:91591559; PID:91591563; TIGR:MG0885; PI
 C;Genetics:
 C;Function: <DP>
 A;Start codon: ATG
 C;Function: <EN1>
 A;Description: as DNA endonuclease PI-M1aI, catalyzes the hydrolysis of internal phosphates
 C;Function: <EN2>
 A;Description: as DNA endonuclease PI-M1aII, catalyzes the hydrolysis of internal phosphates
 C;Superfamily: DNA-directed DNA polymerase KOD
 C;Keywords: endonuclease; hydrolase; nucleotidyl transferase; protein splicing
 Db 182 KODIVIVSKTNDDDEI-----LTGPERRKTSFVYKLISYDPIIVGYNDA 230
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 122 IDKDLACMGIAEKEGKOGGVIRSYKTERVERIPRMEPELKLMLVDFCEMLSSFGMPEPE 181
 QY 154 KGMIMISYADEEARVITWNKNDLPHYDWSMWEREMKRFQIVREKDPDVITYNDN 213
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 182 KODIVIVSKTNDDDEI-----LTGPERRKTSFVYKLISYDPIIVGYNDA 230
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 214 FDPYPLIKRRAEKIGVTLGDRMEHPEPKIRNGDSFAVEIKRGRHEDLFPVVRRTLP 273
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 231 FDPYPLIKRRAEKIGVTLGDRMEHPEPKIRNGDSFAVEIKRGRHEDLFPVVRRTLP 273
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 QY 274 TYTLEAVYEAFLGKTK-SKLGAFEEIAIWETEESMMLQAOYSNEDARATEYELGKEEPM 332
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 285 IKKLENVAAEFLGLKIELDIAKDIYRWSRKE-EKVLNAYRQDAINTVYLAKELLPMH 343
 QY 333 AELAKLIGQSWVDSRSSTGNLWVLYLIRVAVERNELAPKDDDEERRRLRTYLGGIV 392
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Query Match 32.1%; Score 1291; DB 2; Length 1634;
 Best Local Similarity 22.7%; Pred. No. 1.6e-58;

Db 344 YELSMIRLPVDDVTRMGRGKQDVNLSSBAKKGEIAPNPPEAE-----SVEGFV 396
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 QY 393 KEPERGLWNTYIQLRCYLPSIWTNNSPDTLREGCERN-YVAPIVGKCKDFG 450
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 397 LEPERGLHVNWACLDFASWMPSIMAFNISPDTY--GCRCDCYB-APEVGHKFRSPDG 452
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 451 FIPSLIGELITMRQBIKKW-ATIDPIEKKMDYQRQRAVHLHANSYGSYMPKARWTS 509
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 453 FFKRFLRMJLEKREKLVEKLNSPESSEYKFLDQOTLKVLTNSFGYMGWNLARWC 512
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 QY 510 KECAESVWAWGRHYTEMTEKEERFGKYLADDFGFWATPGEKPKETIKKAKEFLY 569
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 513 HPCAETTANGRHFRTSAK-IAESMGFKLYGIDDSIFVTKAGMKTEKNDR---LIK 567
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 570 INSKPGLLMEYSGFYLGFVAKRVLAAVGKRPGVTSIVLKGSGRIGDRAPEDEQ 629
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 568 LHEELP--IQLIEVDEYYSATFFVVERKRYAGLTDEDGRLVKGLEYVRGDDWCLAKKQRE 625
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 QY 630 LEAIKEDSVKEAVIYKVWEEAKYQPLEKLYHQITKDSEYKAIGPHVIAKRL 689
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 626 TEVILKEKIPKEPAISLKVQDVLTRIKEGVSLWVUHQKSLPSKYESMOAHVKAALK 685
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 686 REMGIYPPSSKGIVVIVKGSNIGDRAYPAPIDLEDFDSENLRKTKSIEKKLDKVY 745
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 QY 735 IENQVLPVYLAFGKRYKEDLKYQSSKQVGLAW 771
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 746 IDNQIIPSVPRLERFGYIPEASLK-GSSQMSLDSF 779
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 RESULT 9
 E64410 DNA-directed DNA polymerase (EC 2.7.7.7) family B. Intein containing precursor - Methanococcus jannaschii
 C;Species: Methanococcus jannaschii
 C;Accession: E64410
 C;Date: 13-Sep-1995 #sequence_revision 13-Sep-1996 #text_change 13-Sep-1998
 R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Merrick, J.M.; Glodek, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Blakerson, J.D.; Sodow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.; Science 273, 1058-1073, 1996
 A;Authors: Kaine, B.P.; Borodovsky, M.; Kleck, H.P.; Fraser, C.M.; Smith, H.O.; Woese C;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
 A;Reference number: A64300; MUID:9633799
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-1634 <BUL>
 A;Cross-references: GB:U67532; GB:L77117; NID:91591559; PID:91591563; TIGR:MG0885; PI
 A;Map Position: REW816304-811400
 A;Genetics:
 A;Function: <DP>
 A;Description: as DNA-directed DNA polymerase, catalyzes the polymerization of DNA at C;Function: <EN1>
 A;Description: as DNA endonuclease PI-M1aI, catalyzes the hydrolysis of internal phosphates
 C;Function: <EN2>
 A;Description: as DNA endonuclease PI-M1aII, catalyzes the hydrolysis of internal phosphates
 C;Superfamily: DNA-directed DNA polymerase KOD
 C;Keywords: endonuclease; hydrolase; nucleotidyl transferase; protein splicing
 Db 1 F1;425,795-882,1357-1364/Product: DNA-directed DNA polymerase family B #status predicted
 F;425-795/Cross-link: peptide (Arg-Ser) #status predicted <XTR>
 F;426-794/Product: DNA endonuclease PI-I (pol B Intein 1) #status predicted <WAT2>
 F;883-1358/Domain: DNA-directed DNA polymerase family B extein 2 #status predicted <XT>
 F;883-1358/Domain: DNA endonuclease PI-II (pol B Intein 2) #status predicted <MAT3>
 F;1358-1634/Domain: DNA-directed DNA polymerase family B extein 3 #status predicted <XT>
 F;882-1359/Cross-link: peptide (Arg-Ser) #status predicted
 F;882-1359/Cross-link: peptide (Asn-Ser) #status predicted

Matches	Conservative	Mismatches	Indels	Gaps
Qy	2 IFDTDYITDKCKPPIRIFKENGKENGKIELDPHFFQPYIYALLK---DSDAIDEK--ATK	55		Qy 494
Db	13 LIDNTYKTIKEDKAVIYLYLINS---ILKDRDFKPYFVHLKKEVENEDEIEKIKEFLK	68		Db 1076 LIDNTYKINKVVDIKILDEEKFKAYKVKIDTVIHBGNFKKGRCNIQYIKDVIANIPD
Qy	56 GERRKIVRUVDAVKVKK-FLGRDVEWKLIFERPDVPLRGKIREHPAVIDYEDI	114		Qy 494
Db	69 ND---LLKVENIEVKVITLREKEVITAHPOVKLR-KIKECETVKEIYEDI	123		Db 1136 KEFEDCEIGASGKINALLKDEKLAKFLGFFVTRGLKKQKLGETVYISVYKSLPEV
Qy	115 PFAKYLIDKGLPM---EGDE---ELKUMAFDTEFWIEGDERGKG-E-TIMS	161		Qy 494
Db	124 PFAKYLIDNEIPMRYWDFENPKVSIETPKLKSVAFDMEVYINRDETPNPERDILMAS	183		Db 1196 QKEIAETFKEVFGAGSMVKVKTMDNKIVVZLVLKVIFKCCDDKDHPIELFLASESVIK
Qy	162 YADEEARVITWKNIDLPYDVVSNSNEMRKFVOLVREKDPDVLITYNGDNFDLPLYLIK	221		Qy 494
Db	184 FWDENGKVYTYKEFHNPHNTEVVKNEKELIKLIELTKEV-DVITYNGDNFPYPLKA	241		Db 1256 SFLDGELAKKNSHKGTSFMAKDERYLNQIMILNLUVGIPTRTPVKNQGYKLTNPVY
Qy	222 RAEKLGVTLIGDRKEHPEKIKHMGDSFAVEIKGRIHFLDFPVYRTNLPTTLEAVY	281		Qy 494
Db	242 RAKIVGIDINNGKDFE-ELKIKRGMEYRSYIYGRVHIDLPLYSRRLIKLTKEVTLDDV	299		Db 1316 GTVKDMLDEWKEIEAFAYSGYVYDLSVEDENENFLVNIAHNNSVYGLAPRARRYSE
Qy	282 EAVLGKTKSNAEAAIAWETEESMMLAQYSMDARATYELGKREFFPEAEALAKLIGO	341		Qy 512 CAESVPAWGVYIEMTKEKEFVKLADTDRYATPGK--PETTKKAKFELK
Db	300 YNLFGIEKLKIPHTKLVDYWNND--KTLIEYSLQDAKYKVFPELVMFSRIVNO	357		Db 1376 CAEIVYLYGRVYLEVKE-AEKFGKVILIDTDEYFA-TWKEKISKEELJKKAMEFVY
Qy	342 SWDWDVSRSSTGNLWVWLLRAYERNELAPNPKDDEEYRRRLRTYLGGVKEPERGLWE	401		Qy 570 INSKUPLGLELEVEGFYLREFVAKKRYAVIDEGRITLGLEVRRDMSAEIATEQAVY
Db	358 TPEELTRMSGQMVWLLMRAFKANMIVNPDKPDEEYRRVLTTEGGVKEPEKGME	417		Db 1434 INSKLPGTMELFEGYFVKGJFVFTKRYALIDENGRTVKGLEFVRDWSNTAKITQRRV
Qy	402 NITYLDFRC-----	410		Qy 630 LEAIKEDSEKAVEVKDVEEIAKIQVDELEKLYIHEQITKDSSEYKATGPHVAIKR
Db	418 DIISADFRCPKPKVWVKKGIGIWINEDKEGNYVVLGIDOWQVKVVKWKEYESELINV	477		Db 1494 LEALUVLEGSTEAKKIQDVKLDRBKIKKEDLIYIQTJTKPKEYKTTAPHVEIJKU
Qy	411 -----	410		Qy 690 AAKGIVKVRPSTISYVLRGSGKISDRVILSEYDPKKKYDPDYIENQVPAVLRLTE
Db	478 GLKCTPNHKIPLRYKIKKKINKNDLVRODIYAKSLLTRKGERKLICDFETIGNYK	537		Qy 1554 MREGRKIVGSDIGIIVLKSTKSISERAKI-PEEVD-IDDIVNYIDNOILPPLRIME
Qy	411 -----	410		Qy 750 AFGYRKEDLKQSSKQVGLHAWLK
Db	538 YINIDMDEDFTLKSLELIGLLRRDIEFPSSRGKKRISHQYRVEITVNDEKDFI	597		Db 1612 AVGVSKNELKKEGA-OLTLDKFFK
Qy	411 -----	410		RESULT 10
Db	598 EKIKYIEKKLENVELYVRKRGKTAITLGCAKKIYKIEEIKNKEVYLPNA1LRGFE	657		C69028
Qy	411 -----	410		DNA-dependent DNA polymerase family B (polB) - Methanobacterium thermoautotrophicum
Db	658 GDCVNTVRAVWVQNTNWDKIKFELASLDRGKIKSYTIVYERGKLLKRWIEF	717		C;Species: Methanobacterium thermoautotrophicum
Db	411 -----	410		C;Date: 03-Dec-1997 #sequence-revision 05-Dec-1997 #text_change 22-Oct-1999
Qy	718 SKGDLLKFSLISITSRKRNLLNFTIORTLYKIGDYFYDLDVVCVSLESYKGEVYDL	777		C;Accession: C69028
Qy	411 -----LYPSILTHANSPDTLREGCKNVDIA-PIVYKEKDFPGF	451		R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, R.; Oiu, D.; Spadacina, R.; Viccire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noedding, J.; Reeve, J.N.; Bacteriol, 179, 7135-7155, 1997
Db	778 TLEGRPYYFANGILTHNSVPSIISYNTSPDTLCECCK--DSEKLGHWRCKKKGEGL	835		A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: f4
Qy	452 IPSIIGELITMRQETKKMK---ATDPI--EKKMDYDORAVKHA-----	493		A;Reference number: A69000; MNUID:98037514
Db	836 IPKTLRNLERRRINKRKKMKAEGEINEBYNLUDEOKSLKTLANSITPDEVITIEE	895		A;Accession: C69028
Qy	494 -----	493		A;Status: preliminary; nucleic acid sequence not shown; translation not shown
Db	896 DGIVKVKIGEVYDIDLMRKHKDKIKFSGISILETKNLKTFSDKITKCEIKVKALIRH	955		A;Molecule type: DNA
Qy	494 -----	493		A;Residues: 1-586 <ATH>
Db	956 PYFGKAYKIKLRSRGTIVTRGHSLFKYENGKIVEVKGDDVYREGDOLIVVPKLPCVDKEV	1015		A;Cross-references: GB:AE000888; GB:AE00666; NID:92622304; PIDN:AAB85697.1; PID:926; C;Experimental source: strain Delta H
Qy	494 -----	493		C;Genetics: MTH1208
Db	1016 VINTPKRLINADEEIKDVKVVKWFLKKTLEDIENNKVWIDDCILYKELG	1075		Query Match 30.0%; Score 1209.5; DB 2; Length 586; Best Local Similarity 42.9%; Pred. No. 6.4e-55; Matches 255; Conservative 107; Mismatches 176; Indels 57; Gaps 10;

Db 67 EK-----LEVKEMDLGRPEVIRIEFRHQDPVKIRDRIDLESVRDIREHDIP 116
 Qy 116 FAKRYLIDKGKIPRE-----GDEELKUMAEDI 142
 Db 117 FYRRLIDKISIVPMEELERQGVVEUDSAPSVTIDVTYRVEVTGRYQSTGSAHGLDLSFDI 176
 Qy 143 ETFYHEG-DEFGKEIIMSYA-DEEARVITWNKNIDLPYDVVSNEREMIKRFVQIREKD 200
 Db 177 EVRUPHGMPEPKEDIEVMGAGNGYESVISTAGDHLDIFEVVEDEERELERAEIVID 236
 Qy 201 KDPDVITYNGDNDPLYPLIKRAEKGVLGVTILLGRDKEHEPEKPLHMGDSFAVEIKGRIFDL 258
 Db 237 KKPDLILVGNSDNFDPYTRRAILGAELDGLNGDS--KRTMRGFANATAIKGY 292
 Qy 259 HFDLFPVYRPTINIPPTYEAIVAVGKTSKIGAEEAATIETEESUKKLAQYSMEDA 318
 Db 293 HVQDIPVMPYRMMYDRLTYRVEOFLGSEKIDPGRDWEYMDRDELDRFLFRRSY 352
 Qy 319 RATYELGKEFPMEALAKLICQSVWDSRSSTGNLVEYLRAVERNELAPNKDEEE 378
 Db 353 VATHRIAEKILPLNIELTULVQGQPLFDISRMATGQQAEMFLVRYKAYQGELVPLKPSQD 412
 Qy 379 YRRRLRTTYLGYYKEPERGLWENITYLDFRCIPLXPSITWHTNTSPDTLREGCKNYDAP 438
 Db 413 FSSRRGRRAVGGYKPERGLHENIQFDRSLYPSIISKNSPDLTDDSECYVAP 472
 Qy 439 IVGKFKCOPGPFGPSIISLGEILINRQEKKMATIDPIEKKMDYRORAVKLHANSYY 498
 Db 473 EYGYRKFRKSPRGRVPSVIGELSRVREKEMGDDPERKLINVQDALKLANTMYG 532
 Qy 499 YMCPKPARWYSKCAGCAESVTAWGRHYIEMTIKEEKGKVLYADTDGFYATPG 553
 Db 533 VGYSRFRWYSMCEAEATPAWGRDYIKKTTAEE-FGFHTVADTGEYATVYRG 586
 RESULT 11
 JC382
 DNA-directed DNA polymerase (EC 2.7.7.7) B3 - Sulfurisphaera ohwakuensis
 N; Alternative names: DNA polymerase alpha, beta, gamma; DNA polymerase I, II, III
 C; Species: Sulfurisphaera ohwakuensis
 C; Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 31-Dec-2000
 C; Accession: JC7382
 R; Iwai, T.; Kuroswa, N.; Itoh, Y.H.; Kimura, N.; Horiochi, T.
 DNA Res. 7, 243-251, 2000
 A; Title: Sequence analysis of three family B DNA polymerases from the thermoacidophilic
 A; Accession: JC7382
 A; Molecule type: DNA
 A; Residues: 1-781 <IWA>
 A; Cross-references: DDBJ:AB032375
 C; Comment: this enzyme has both 3',-5' exonuclease and polymerase activities, and plays
 A; Gene: B3
 C; Keywords: DNA replication; exonuclease; metal binding; nucleotidyltransferase
 RESULT 12
 B56277
 DNA-directed DNA polymerase (EC 2.7.7.7) II - Pyrodictium occultum
 C; Species: Pyrodictium occultum
 C; Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 20-Jun-2000
 C; Accession: B56277
 R; Uemori, T.; Ishino, Y.; Doi, H.; Kato, I.
 J. Bacteriol. 177, 2164-2177, 1995
 A; Title: The hyperthermophilic archaeon Pyrodictium occultum has two alpha-like DNA P
 A; Reference number: A56277; MUID:95238290
 A; Accession: B56277
 A; Molecule type: DNA
 A; Status: preliminary
 A; Molecule type: DNA
 A; Cross-references: GB:D38574; NID:9807829; PIDN:BAA07580.1; PID:9807830
 C; Superfamily: herpesvirus DNA-directed DNA Polymerase
 C; Keywords: nucleotidyltransferase
 Query Match 29.9%; Score 1204.5; DB 2; Length 781;
 Best Local Similarity 36.6%; Pred. No. 1.7e-54; Mismatches 281; Conservative 164; Gaps 21;
 Matches 204; Indels 87; Gaps 21;
 Qy 2 IFFDYLTKDGKPIRERK-KENGEFKIELDPHQPYTVALLDSAIDEIK-AIKGER 58
 Db 7 ILDPSYDVVNEVKPVYIWWIDKSEGVRVVLLEKKFRPYFYALVVDSYNIDEIRKTLKSL 65
 Qy 59 HGGKIVRVRDAVKKKFGLRDVWKLLEHPDQPALRGKIREHPAVIDYEDIPAK 118
 Db 67 PYSPITSIDV--EEKYFGSPVVKLKIETVPAVYRVRDEVAIKGYSVLADRIFYM 124
 Qy 119 RYLDKGKIPM-----EGE-----EIKLMDIETYH 147
 Db 125 RYSDINUKPFWIEAEVEEIKENNFRYKVKVYELKINKYKPELKVLADEVY- 182
 Qy 148 EGDERG---KGEIMISYADEEERAVITWNKNIDLPYDVVSNEREMIKRFVQIREKD 202
 Db 183 --NKYSPNPDRDPVITGWTKEGKQFLADKV-----DLRAIREFINVQTYD 231
 Qy 203 PDVLVITYNGDNDPLYPLIKRAEKGVLGVTILLGRDKEHEPEKPLHMGDSFAVEIKGRIFDL 262
 Db 232 PDTIVGYNINNFDWMVYLLERANIRGIRLDVGR-RVNGEPS--QGVYGHYSITGRNLVDL 287
 Qy 263 FPWVTRINIPPTYEAIVAVGKTSKIGA-EELAIIWETEESMKKLAQYSMEDA 319
 Db 288 YGFAOSIQEVVKVTLENIDYLGVLPEKEKRTIVEWYDIPKQWYDDEKKRDILKNUDDAK 347
 Qy 320 ATYELGKEFPMEALAKLICQSVWDSRSSTGNLVEYLRAVERNELAPNKDEEE 379
 Db 348 SAYLGEVIFPGLTTRISGLPQDLSMASYGVHWEVLMREAVKNEJLPK-EFERY 406
 Qy 380 RRLRLRTTYLGYYKEPERGLWENITYLDFRCIPLXPSITWHTNTSPDTLREGCKNYDAP 439
 Db 407 E---SYEGGLVNSPLPGTHEEVVLFDSMSYPSIMKVNIGPTLVKGECECWSP- 460
 Qy 440 VGYFKCKDFGFPFISITLGEJLTMROBIEKIKMKATIDPIEKKMDYRORAVKLHANSYY 499
 Db 461 VGHKRKEPPGPLYKVNLEKLIQERKEVKVLMKETIDDEYDKVLDARQRAKVMANAFSY 520
 Qy 500 MGYPKARWYSKCAGCAESVTAWGRHYIEMTIKEEKGKVLYADTDGFYATPG 559
 Db 521 MGWLGRWYSGEKAATVANTQRQITSDASAKI-TAKEG-GFTWVYGDTSIVVGGGD---- 574
 Qy 560 KKAKAEFLKVINSKNPLGLLELEYFYLGFVVA-KRRAVIADEGRITRGLVRRD 618
 Db 575 ---INSLITISSKG--LEIKIDKIVKRFVFTENKKRAGLTDGKIDVGFENVGRW 629
 Qy 619 SETAKETQAKVLEATIKEDSVKAIVEIVKDVVEEAKYQVPLEKLVTHOITKOLESEYKA 678
 Db 630 CDLAKQVOTVNLIELIKSGKVEDAIKYKSVIFDLYRNFRIEDLILWIKYIDKNLDEYD 689
 Qy 679 1GPHVIAKQKLAQKTKVRCGTLISYIVLARGSKISDRV--ILSEYDPEKKVYDPPY 735
 Db 690 TAPHVVAKKAKAGSYLVSKVGKTYVIVKGSKISDKAEPYFLVKE---KNKIDVEY 745
 Qy 736 IENQLPVAVRILEAFGYREDLK 759
 Db 746 IDKQJIPVALRILEFGFGVKBSSLK 769
 RESULT 13
 B56277
 DNA-directed DNA polymerase (EC 2.7.7.7) II - Pyrodictium occultum
 C; Species: Pyrodictium occultum
 C; Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 20-Jun-2000
 C; Accession: B56277
 R; Uemori, T.; Ishino, Y.; Doi, H.; Kato, I.
 J. Bacteriol. 177, 2164-2177, 1995
 A; Title: The hyperthermophilic archaeon Pyrodictium occultum has two alpha-like DNA P
 A; Reference number: A56277; MUID:95238290
 A; Accession: B56277
 A; Molecule type: DNA
 A; Status: preliminary
 A; Molecule type: DNA
 A; Cross-references: GB:D38574; NID:9807829; PIDN:BAA07580.1; PID:9807830
 C; Superfamily: herpesvirus DNA-directed DNA Polymerase
 C; Keywords: nucleotidyltransferase
 Query Match 29.6%; Score 1190; DB 2; Length 803;
 Best Local Similarity 33.7%; Pred. No. 9.8e-54; Mismatches 278; Indels 110; Gaps 23;
 Matches 281; Conservative 164; Mismatches 278; Indels 110; Gaps 23;
 Qy 1 MIFDDY-----ITKDGKPIRERK-KENGEFKIELDPHQPYTVALLDSAIDEIK-AIKGER 45
 Db 8 VLLDSSYEVILGKEPVYIWWIDKSEGVRVVLLEKKFRPYFYALVVDSYNIDEIRKTLKSL 55
 Qy 46 SAIDEIKAKGERRHGVIR--VVDAVKKKFLGLRDVWKLFFEHPODPALRGKIREH 103

Query Match 28.3% Score 1138; DB 2; Length 784;
 Best Local Similarity 34.6%; Pred. No. 4 5e-51;
 Matches 282; Conservative 137; Mismatches 287; Indels 110; Gaps 19;

Db 56 DMVEETAA--SIRLSVVKSPIIDAKPLDKRYFGRKAVKITMIPESVRYHREAVKKI 113
 Qy 104 PAVIDIYEVDPFAXYRLIDKGL-----IPME-----GDEE----- 134
 Db 114 EGVEERSLEADDFRFAFMYLYDOKRKYFTVYRIPVEDAGRNFGFRVDRVYKAGDPEPLADI 173
 Qy 135 -----LKLMAFDIETFYHGD-BFGKGETIMISYADEE-EARVITWKIDLPKVDVNS 185
 Db 174 TRIDLPPMRLVAFDVEVYRSRGSPNPARDPDTIVSILRSDBGKERLIEAEGHD----- 225
 Qy 186 NEREMTIRKFVQIVREKDPDVILTYNGDNFDLPLYLIKRAEKGVTLLGRDKEHPEPKIHR 245
 Db 226 -DVRVLFREFFYVRAFDPTDITVGYNSNHFDPMYLMEARRLGKIDVTR-RVGAEPPTSV 283
 Qy 246 MGDSFAVEIKGRIHDPLPVYRRTINLPTTYLEAVYEVIGKTAKKGAEE---TAIWE 302
 Db 284 YGH---VSVGRNLNDYDVAEEMBPIKMTLEAVAYLGKMKSERVIEWWRIPEWYD 340
 Qy 303 TEESMKKLAQYSMEDARATYELGKEFPMAELAKLIGOSWVDSRSSTGNLVEYLRLV 362
 Db 341 DEKKRQOLLERYALDDVYRATYGLAEKMLPFAIQLSTVGTGVLDQVAMGVGRLENYLMLRA 400
 Qy 363 AYERMLAPNPKPDEEFYRRLRTTIGGYKKEPERGLWENITYLDFRCLYPSITVHNVS 422
 Db 401 AYDMNELLVPRVE-----RRGEGSYKGVAVLKLPLKGVHENWVLDFFSMWPSIMKYNG 454
 Qy 423 PDTL--ERECKNVD---VAPIVGKFCDFGPFTSPLSILBLITMROEIKKKMA-TIDP 476
 Db 455 PDTIYDPSCECPKYGCCYVAPEGHFRRSPPGFFKTVLENLKLQRQYEMKEPPDS 514
 Qy 477 TEKKMLDYORAVKLUHANSVYGYMGPKARWYSKRCAESVYAWGRHYIEMPTKEEKG 536
 Db 515 PEYRUYDEROKALKVLANASYGGMWSHARWYCKRCAEVTAWGRNL-LTAIEYARKLG 573
 Qy 537 FKVLYKDGTGFTATPGKEPKETIKKAKELKYLKNSKPLLELYEGFLRGGFV-AKK 595
 Db 574 LKVYGDTSFLFWVVD-----KEKVKELTVEKRELG--FEIKLDKTYKVFTEAKK 624
 Qy 595 RYAVIDEERGTRTGLEVVRDWWSETAKETQAKLVEALKEDSVKAVETVKDVEEIAK 655
 Db 625 RYVGLLEDGRIDIVFEAVRGDWCELAKEQEAKEELVWNGNWKAISTIREVKOLRE 684
 Qy 656 YOVPLEKLVTHEQTKRDSEYKATGPHVATAKRLAAKGKIVKVRGHTISYVLRGSGKISD 715
 Db 685 GKVPIPTKLIWTKSRKIEEYEHADPHVMARRKMEAGYEVSPGDKVYVIVKGSVSS 744
 Qy 716 RVLILSEYDPKKKHPDYYENQVLPAVRLFEATGYRKEDLKQSSKVG 768
 Db 745 RAYPYFMVDPSS-TIDVNYVTDHQIVPAALRLSYFGVTEROKLAATVORSL 795
 .Db -

RESULT 13

E7.2515 probable DNA-directed DNA polymerase APE2098 - Aeropyrum pernix (strain K1)
 C;Species: Aeropyrum pernix
 C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
 C;Accession: E7.2515
 C;Kavarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; KDNA Res. 6, 83-101, 1999
 A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
 A;Reference number: A72450; NID:99310339
 A;Accession: E7.2515
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-784 <XW>
 A;Cross-references: DDBJ:AP000063; NID:95105654; PIDN:BAA81109.1; PID:95105797
 A;Experimental source: strain K1
 C;Genetics:
 A;Gene: APE2098
 C;Superfamily: herpesvirus DNA-directed DNA polymerase

Query Match 28.3% Score 1138; DB 2; Length 784;
 Best Local Similarity 34.6%; Pred. No. 4 5e-51;
 Matches 282; Conservative 137; Mismatches 287; Indels 110; Gaps 19;

Db 11 DOKPITIRIFKENGEEFKIELDPHFQPYTIALIKDDSAIDEIKATIGERHGKIVYDAVK 70
 Qy 16 DGSRRVVFY---GEFR-----PYFYVLPDGSVGLDQAMIRRLSRSPSPILVER 63
 Db 71 VKKFELGRDVEWVKLUFEEHODPVPLRGKLEHHPAVIDYDIPAKRVLIDKGLPM- 129
 Qy 64 VRRFEGREVERALKVTLVLPASVREYRAVRLIGGVDLADIPALRFIDENYPMR 123
 Qy 130 -----EGD-----EFLKMAFDIETFYH-EGDEF 152
 Db 124 WVAEVREVAVPHGSYVDRATYLSGDIREDETRIOEDPLKGRLVMAFDIEVYSKMRTPD 183
 Qy 153 GKEGIMISYAD-----EEARVITWKNTDLPVVDVNSNERENIKRKFVQIVREKDP 204
 Db 184 KKDPPVIMIGLQAGGFEIILRADM-----SDKVIAGFVERVSKSDP 227
 Qy 205 VLTIVYNGDFDPLYKLRKAELGVTLLGRDKEHPEPKIHMGDSFAVELKGRIHDLP 264
 Db 228 VIVGYNONRFDPYLMVERAVRLGVKLAVGRESVEPGLG----HYSGRGLNVDL 282
 Qy 321 TYELGKEFPMAELAKLIGOSWVDSRSSTGNLVEYLRLAYERNELAKNPKDDEYR 380
 Db 342 TMGLAKFPLGAELQSGLPDLQDMAASVGRFLWRLREAKLGEYLVPNVERSE 401
 Qy 381 RLRITYLGVYKEPERGLWENITYLDFRCYPSITVHNVSPTDLECGKND----- 435
 Db 402 -----YAGATVLRPREGHEDIAVLDFAFSMWNPKNWKYNGPDTLVRGS-EEYGBEEVY 454
 Qy 436 VAPIVYKEKDFGPFTSPLSILBLITMROEIKKKMA-TIDPTEKKMDYFRAVKLHN 494
 Db 455 TAPEGHKFRKSPPGFFKKLTERFLSWRQIRSEMMKKHPPDSPEYKLLEROKATKLAN 514
 Qy 495 SYVGGMGKPKARWYSCEAESTANGRYHETMTKIEEKGFKVYADPGFYATPG 554
 Db 515 ASYGMGMWPHARWYCRBCEAVTAWGRSRITPAIRAGE-LGLEVYGDITSLFVNDP 573
 Qy 555 KPETIKKAKELKYLKNSKPLGLELEYEGFLRGGFV-AKKRAYVIDEGRITRGL 613
 Db 574 KVERLIRFVEELGF-----DIKVYKVRVFTEEKRYVGLTVOGKIDVGEA 624
 Qy 614 VRDWSBIAKETQAKYLEALKEDSVKAVETVKDVEEIAKQVPLEKLYTHEQTKDL 673
 Db 625 VRGDSWSELAKETOFKYAEVLTGTSYDEAVDVRNTEKLRQGDVMRKVWIKLJRP 684
 Qy 674 SEYKAIGGPHVIAKRAAKGKIVRGPETISYVLRGSGKISDRVILSEYDPKKHYDP 733
 Db 685 SMYEAROPHTAHLIMERAGKIVVERGAKGYVVTKGSPLYTRA--KPYFMSASKEEVDE 742
 Qy 734 YYENQVLPAVRLERAFGKEDLKQYQSKQVGLD 769
 Db 743 YYVDQKQVPAALRLIQYFGVYKEKLKGGRSTL 778
 .Db -

RESULT 14

T05731 DNA-directed DNA polymerase (EC 2.7.7.7) delta chain - soybean
 C;Species: Glycine max (soybean)
 C;Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 18-Jun-1999
 C;Accession: E7.2511
 R;Collins, J. T. B.; Cannon, G. C.; Heinhorst, S
 Submitted to the EMBL Data Library, August 1997
 A;Reference number: 215439
 A;Accession: T05731
 A;Status: preliminary; translated from GB/EMBL/DDBJ

A: Molecule type: mRNA	A: Residues: 1-1088 <COG>
A: Cross-references: EMBL:AF020193; NID:92895197; PIDN: AAC13443.1; PID:92895198	C: Genetics
C: Function: Pol delta	A: Gene: Pol delta
A: Description: catalyzes replication of DNA	C: Superfamily: herpesvirus DNA-directed DNA polymerase
C: Keywords: DNA binding; DNA replication; nucleotidyltransferase	
Query Match 17.2%; Score 693.5; DB 2; Length 1088; Best Local Similarity 27.2%; Pred. No. 5e-28; Indels 161; Gaps 30	Matches 237; Conservative 149; Mismatches 324; Indels 161; Gaps 30
QY 15 IIRIFKENGKFKELDH-FQPYIYALIKKDDSAIDEI-----KAIGSERHK 61	QY 15 IIRIFKENGKFKELDH-FQPYIYALIKKDDSAIDEI-----KAIGSERHK 61
Db 88 IIRIFGVTKEGHHSVCCNTHGFPFYICCPPGMGPDIDSHFHQTLEGMRMEANRNSVKG 147	Db 88 IIRIFGVTKEGHHSVCCNTHGFPFYICCPPGMGPDIDSHFHQTLEGMRMEANRNSVKG 147
QY 62 IIRVNDAYRKV-----KTFGLGRDVEWKLIFEHTQDVPAURGKTRBPAVD-- 108	QY 62 IIRVNDAYRKV-----KTFGLGRDVEWKLIFEHTQDVPAURGKTRBPAVD-- 108
Db 148 FVRIIEMVQRRTMIVYQOSSNSQPL-----KIVVALPTMVASCRG-ILDRGIGLDLG 199	Db 148 FVRIIEMVQRRTMIVYQOSSNSQPL-----KIVVALPTMVASCRG-ILDRGIGLDLG 199
QY 109 -----IEYDIPPAKRYLIDKGLI----- 127	QY 109 -----IEYDIPPAKRYLIDKGLI----- 127
Db 200 MKSFLTYESNVLFALRFMDICNTVGGNNWIGIPAKYKKRTAKSLSYCQLEFDCLYSLELISH 259	Db 200 MKSFLTYESNVLFALRFMDICNTVGGNNWIGIPAKYKKRTAKSLSYCQLEFDCLYSLELISH 259
QY 128 PMEGD---EELKLMADFETTFYHEG - DDFGKGEBIIMSYADEEAEVITWKNIDLPY- 180	QY 128 PMEGD---EELKLMADFETTFYHEG - DDFGKGEBIIMSYADEEAEVITWKNIDLPY- 180
Db 260 APEGEYSKMAPPRLSFDIECAGRKGHFREPTHPVQI-----ANLYTQGEDQPF 312	Db 260 APEGEYSKMAPPRLSFDIECAGRKGHFREPTHPVQI-----ANLYTQGEDQPF 312
QY 181 -----VDVW---SNEEREMIKFVQTYREKDDPLVITYNGNDNFDLPLYIKRA 224	QY 181 -----VDVW---SNEEREMIKFVQTYREKDDPLVITYNGNDNFDLPLYIKRA 224
Db 313 RNVMTLKSCTPITGVDMVMPETEERVILAWRDFIREVDPDIIIGNICKFDLPLYIILER 372	Db 313 RNVMTLKSCTPITGVDMVMPETEERVILAWRDFIREVDPDIIIGNICKFDLPLYIILER 372
QY 225 KLGVT---LLGRKREHP-----EPKTHRMGGSFAVEKGRHIFDLPFLPVURTRNLP 274	QY 225 KLGVT---LLGRKREHP-----EPKTHRMGGSFAVEKGRHIFDLPFLPVURTRNLP 274
Db 373 NLIKIAEFPILGIRRNRSVYRVKDTFSSRQYGTRESKEAVAVEGRVYFDLQLQYMQDVKLSS 432	Db 373 NLIKIAEFPILGIRRNRSVYRVKDTFSSRQYGTRESKEAVAVEGRVYFDLQLQYMQDVKLSS 432
QY 275 YTEAVVYAVLGRKTSKLGAAETTAIWE-TEESMKKLAQSMEDARATEYLGKE-FFPM 331	QY 275 YTEAVVYAVLGRKTSKLGAAETTAIWE-TEESMKKLAQSMEDARATEYLGKE-FFPM 331
Db 433 YSLNSVSSHFLSFOKEVDHHSIISDLQNGNAETRRRLAYVCLKDRYLPQRLDQLKLMFIYN 492	Db 433 YSLNSVSSHFLSFOKEVDHHSIISDLQNGNAETRRRLAYVCLKDRYLPQRLDQLKLMFIYN 492
QY 332 EASLAKLKGQSV-WDVRSRSTGNLVEWVYLIVRATERNELAPN-KPDEEEXRRLRRTYLG 389	QY 332 EASLAKLKGQSV-WDVRSRSTGNLVEWVYLIVRATERNELAPN-KPDEEEXRRLRRTYLG 389
Db 493 YVEMARVGVPIFLSLSRQOSIKLVS-QLRRAROKNLVTPNAQGSE-----QGTEG 546	Db 493 YVEMARVGVPIFLSLSRQOSIKLVS-QLRRAROKNLVTPNAQGSE-----QGTEG 546
QY 390 GYKKEPERGLWEN-ITYLDFRCPLYSITIWNVSPDTLE-REGCKNYDVA-----IVG 441	QY 390 GYKKEPERGLWEN-ITYLDFRCPLYSITIWNVSPDTLE-REGCKNYDVA-----IVG 441
Db 547 ATWLEARAGFYERKPIATLIFASLWPSIMMAYNLYCUTLWIPEDARKLNIPPSVNRTPSG 606	Db 547 ATWLEARAGFYERKPIATLIFASLWPSIMMAYNLYCUTLWIPEDARKLNIPPSVNRTPSG 606
QY 442 YKFCKD-F-PGFPSIISLGLBETIMRQEIKKRMKATDPTEKKMMDYRQARYKLHANSYYGY 499	QY 442 YKFCKD-F-PGFPSIISLGLBETIMRQEIKKRMKATDPTEKKMMDYRQARYKLHANSYYGY 499
Db 607 ETFVKSNLQKGILPEIIEELLLTARKRRAKADKLERKDPKAVLDGQRQLALKISANSVYGF 666	Db 607 ETFVKSNLQKGILPEIIEELLLTARKRRAKADKLERKDPKAVLDGQRQLALKISANSVYGF 666
QY 500 MGPKPARKMYSKEAESVYAWGRHYIEMTKEIEKF---GF---KVLYADTDSFYATI 551	QY 500 MGPKPARKMYSKEAESVYAWGRHYIEMTKEIEKF---GF---KVLYADTDSFYATI 551
Db 667 TGATIGOLFCLEISSVSYGQMLTKEVYHNAEVITYGDTDSVNVQF 726	Db 667 TGATIGOLFCLEISSVSYGQMLTKEVYHNAEVITYGDTDSVNVQF 726
QY 552 PGKEPKETIKKKAKEFLKINKSINKLPGLELEYEGLYFLRGFVWAKRYAVI----DEEGRI 606	QY 552 PGKEPKETIKKKAKEFLKINKSINKLPGLELEYEGLYFLRGFVWAKRYAVI----DEEGRI 606
Db 727 GYSAAVEEMNLGSEAEHSGTFKPKIPIKEFKEKYYPPVLLISKRYAGLFWTKPNFDKM 786	Db 727 GYSAAVEEMNLGSEAEHSGTFKPKIPIKEFKEKYYPPVLLISKRYAGLFWTKPNFDKM 786
QY 607 TTRGLEVYVFRDWESEIKAETQAKVLEAITLEKEDSVKAVELVKDVWEIAYKQVPLERLVIH 666	QY 607 TTRGLEVYVFRDWESEIKAETQAKVLEAITLEKEDSVKAVELVKDVWEIAYKQVPLERLVIH 666
Db 787 DTKGKETVFRDNCNLVVKNLVNDCLHKILIDRDPGAVQVYKNAISDLNMRNMDLSLVI 846	Db 787 DTKGKETVFRDNCNLVVKNLVNDCLHKILIDRDPGAVQVYKNAISDLNMRNMDLSLVI 846
QY 667 EOIJKDLSYEKAYGPHVATAKRLAAGKIVRP---GIIYIVLRLGSS---GKISDRVI 718	QY 667 EOIJKDLSYEKAYGPHVATAKRLAAGKIVRP---GIIYIVLRLGSS---GKISDRVI 718
Db 847 KGLKTKGTDYEVRAHVELMRKDAATAPNVDGRPVVIIAKGKAYERSEDPY 906	Db 847 KGLKTKGTDYEVRAHVELMRKDAATAPNVDGRPVVIIAKGKAYERSEDPY 906
QY 719 LISEYDPKKKYQDPYYENQVLPDAVLIR 749	QY 719 LISEYDPKKKYQDPYYENQVLPDAVLIR 749

OY	615	RR - - - DWESETAKETOAKVLEAILKEDSVEKADEVYKDVVEEIAKQVQPLEKLVHREQI	669
Db	600	RKGDEEDMLCELARNTKRKTEELILSKDVKAALKVKTIVLRLRGEDFNEELITWAKI	659
OY	670	TKDLSYKAGPHVATAKRLAAGKIVKRPGTISYVLRGSKISDRV - - ILSEDPK	726
Db	660	ERDLNEYNQLPFLVTAARKLAQSIOSGYGLISKDSKIGYVIVKGGLPNDRKEPEFLYKE - -	715
OY	727	KHKYDDYIYIENQVLPVLAVRLLEAFGYRKEDK	759
Db	716	KNRIDEYVQ-FIRFETLKLUKPLGVNEESUK	747

Search completed: May 28, 2002, 14:18:06
Job time: 64 sec

US-08-902-632-2

Query Match 80.1%; Score 3225.5; DB 3; Length 774;
 Best Local Similarity 77.5%; Pred. No. 3; 8e-243; Mismatches 82; Indels 3; Gaps 2;
 Matches 599; Conservative 82; Mismatches 89; Indels 3; Gaps 2;
 Matches 599; Conservative 82; Mismatches 89; Indels 3; Gaps 2;

Qy 1 MIFDTDYITDGKPIRIFKENGKIEKLFDPHQPYIVALLKDDSAIDEKAIGERHG 60
 Db 1 MILDPTYITDGKPIRIFKENGKIEKLFDPHQPYIVALLKDDSAIDEKAIGERHG 60
 Qy 61 KIVRVDVAKVKKFLGRDVEWKLFPHQPYIVALLKDDSAIDEKAIGERHG 60
 Db 61 TVTUVKVERVQKKFLGRDVEWKLFPHQPYIVALLKDDSAIDEKAIGERHG 60
 Qy 121 LIDKGLIPMGEDEELKLMADFOTFYHEGDEGKGGETIMSYADEEAVITWNIDPY 180
 Db 121 LIDKGLIPMGEDEELKLMADFOTFYHEGDEGKGGETIMSYADEEAVITWNIDPY 180
 Qy 181 VDWNSNEMIKRFQVIRKEDPVLITNGDNFDPLYKRAEKGQSYADEGARVITWNIDPY 180
 Db 181 VDWNSNEMIKRFQVIRKEDPVLITNGDNFDPLYKRAEKGQSYADEGARVITWNIDPY 180
 Qy 241 PKIHRGDSAVEIKGRIHDLFPLVRRTINLPTYLEAVYAVLGKTKSGLAEIAI 300
 Db 241 PKIHRGDSAVEIKGRIHDLFPLVRRTINLPTYLEAVYAVLGKTKSGLAEIAI 300
 Db 239 PKIORMGDRFAVEVKGRIHDLFPLVRRTINLPTYLEAVYAVLGKTKSGLAEIAI 298
 Qy 301 WETEESMKKLAQSYMEDARATEVLEGKFFPMEAEALKLIGSWDVRSSTGNTLWYLL 360
 Db 301 WETEESMKKLAQSYMEDARATEVLEGKFFPMEAEALKLIGSWDVRSSTGNTLWYLL 360
 Qy 359 RKAYERNELAPNPKDEKELARR-ROSVEGGYKKEPERGLWENIVLDFRSLYPSIITHN 417
 Db 359 RKAYERNELAPNPKDEKELARR-ROSVEGGYKKEPERGLWENIVLDFRSLYPSIITHN 417
 Qy 421 VSPDTLREREGCKNQYDAPIVYKFCDFPGTIPSITGELITMRDKEKKKATDPIEK 480
 Db 421 VSPDTLREREGCKNQYDAPIVYKFCDFPGTIPSITGELITMRDKEKKKATDPIEK 480
 Qy 481 MLDYRQRAVKLHANSYGYGMYPKARNWYKSCAESTAWGHYIETIKEEKKGFKV 540
 Db 481 MLDYRQRAVKLHANSYGYGMYPKARNWYKSCAESTAWGHYIETIKEEKKGFKV 540
 Qy 478 LLDYRORAIKILANSYGYGYARAWYCKBCAESVTAWGREYIUTIKEEKKGFKV 537
 Db 541 YADTDGYATIPGKEPKETIKKAKEFLYINNSKPGLELEYEFGYLGFPAKARYAVI 600
 Qy 538 YSDTDGFATPGADAEVKKKMELFLYINAKLSCALELEYEGFKRGFVTKKKYAVI 597
 Qy 601 DEEGRITTRGLEVVRDWESETAKETOKVLAIKLDSVERKAVETVYKDVVEEAKYQPL 660
 Db 598 DEEGRITTRGLEVVRDWESETAKETOKVLAIKLDSVERKAVETVYKDVVEEAKYQPL 657
 Qy 661 EKLVIRHQITDSELEYKAIGPHVATKRLANGKIVRPGTISYVLRGSGKISRVILL 720
 Db 658 EKLVIRHQITDSELEYKAIGPHVATKRLANGKIVRPGTISYVLRGSGKISRVILL 717
 Qy 721 SEYDPKKHKKYDYYTENQVLPAVLRLAEGYRKEDLKQVSSKQVGLDAWLK 773
 Db 718 DEFPDKHKYDYYTENQVLPAVLRLAEGYRKEDLKQVSSKQVGLDAWLK 770
 RESULT 2
 US-09-073-354-1

Query Match 80.0%; Score 3222.5; DB 3; Length 774;
 Best Local Similarity 77.8%; Pred. No. 6.5e-243; Mismatches 89; Indels 3; Gaps 2;
 Matches 598; Conservative 83; Mismatches 89; Indels 3; Gaps 2;

Qy 1 MIFDTDYITDGKPIRIFKENGKIEKLFDPHQPYIVALLKDDSAIDEKAIGERHG 60
 Db 1 MILDPTYITDGKPIRIFKENGKIEKLFDPHQPYIVALLKDDSAIDEKAIGERHG 60
 Qy 61 KIVRVDVAKVKKFLGRDVEWKLFPHQPYIVALLKDDSAIDEKAIGERHG 60
 Db 61 TVTUVKVERVQKKFLGRDVEWKLFPHQPYIVALLKDDSAIDEKAIGERHG 60
 Qy 121 LIDKGLIPMGEDEELKLMADFOTFYHEGDEGKGGETIMSYADEEAVITWNIDPY 180
 Db 121 LIDKGLIPMGEDEELKLMADFOTFYHEGDEGKGGETIMSYADEEAVITWNIDPY 180
 Qy 181 VDWNSNEMIKRFQVIRKEDPVLITNGDNFDPLYKRAEKGQSYADEGARVITWNIDPY 180
 Db 181 VDWNSNEMIKRFQVIRKEDPVLITNGDNFDPLYKRAEKGQSYADEGARVITWNIDPY 180
 Qy 241 PKIHRGDSAVEIKGRIHDLFPLVRRTINLPTYLEAVYAVLGKTKSGLAEIAI 300
 Db 239 PKIORMGDRFAVEVKGRIHDLFPLVRRTINLPTYLEAVYAVLGKTKSGLAEIAI 298
 Qy 301 WETEESMKKLAQSYMEDARATEVLEGKFFPMEAEALKLIGSWDVRSSTGNTLWYLL 360
 Db 299 WETEESMKKLAQSYMEDARATEVLEGKFFPMEAEALKLIGSWDVRSSTGNTLWYLL 360
 Qy 361 RVAYERNELAPNPKDEKELARR-ROSVEGGYKKEPERGLWENIVLDFRSLYPSIITHN 420
 Db 359 RKAYERNELAPNPKDEKELARR-ROSVEGGYKKEPERGLWENIVLDFRSLYPSIITHN 417

Db 299 WETGENLERVARYSMEDAKVYELGKEFLPMEAOQLSLIGOSLWDSRSSTGMLNVEWLL 358
 Qy 361 RVAYERNELAPNPKDEEEYRRRLTYLGYWPERGLWENTYLDRCLYSLIVHN 420
 Db 359 RKAYERNELAPNPKDEKELARR-ROSVEGGYVPERGLWENTYLDRSLYSLIITHN 417
 Qy 421 VSPDTLERTGCKKNVAPTYGVYKCFDPGFPIISLGEILTMQOEIKKKMATIDPIEK 480
 Db 418 VSPDTLNREGCKEVDAPYGVHFRCKDPGFISLGLDLEERQIKKKMATIDPIEK 477
 Qy 481 MLDYRQRAVKLHANSYGYGMGPKARWYSKECSESVTAWGRHYTEMTKEIEKFGKV 540
 Db 478 LLDTRQRAKILANSYGYGPKARWYKCECAESVTANGREYITMTKEIEKFGKV 537
 Qy 541 YADDGFYATIPGKRPETIKKKAKEFLYINSKPLGULEYEGFYLGFVAKRYVI 600
 Db 538 YSDTDGFFATIPGADAETVKKAMFLNYINAKLPGALEYEGFYLGFVAKRYVI 597
 Qy 601 DEGRTTGLEVVRDWESEIAKETOAKYLEAIKEDSVAKEVAKIVDVEETAKYQPL 660
 Db 598 DEGKITTGLEVVRDWESEIAKETOARVLEALLKDGVEKAVRKVTEKLSKYEVPP 657
 Qy 661 EKLYTHEQITKDKDLYKATGPHVAKRLLAAG3KVRPOTISIVLRCGSKSDRVL 720
 Db 658 EKLYTHEQITRDLKDYKATGPHVAKRLLAAG3KVRPOTISIVLRCGSKSDRVL 717
 Qy 721 SEYDPKKKHYPDYYIENQVPAVLRLAEGYRKEDLYKQSSKQVGLDAWLK 773
 Db 718 DEFDPTKHKYDAEYIENQVPAVERILRAGFIRKEDLYKQTKRQVGLSAWLK 770

RESULT 6
 US-09-418-027-1
 ; sequence 1, Application US/09418027
 ; Patent No. 622505
 ; GENERAL INFORMATION:
 APPLICANT: KITABAYASHI, Masao
 APPLICANT: ARAKAWA, Taku
 APPLICANT: INOUE, Hiroki
 APPLICANT: KAWAMAMI, Bunsui
 APPLICANT: KAWAMURA, Yoshinisa
 APPLICANT: IMANAKA, Tadayuki
 APPLICANT: TAKAGI, Masahiro
 APPLICANT: MORTIKA, Masaki
 TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Kenyon & Kenyon
 STREET: 1025 Connecticut Avenue, N.W., Suite 600
 CITY: Washington
 STATE: D.C.
 ZIP: 20036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WordPerfect 6.1 Windows
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/418,027
 FILING DATE: 31 MAY 1995
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/073,354
 FILING DATE:
 APPLICATION NUMBER: JP 1404096/95
 FILING DATE: 31 MAY 1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Toffanetti, Judith L.
 REGISTRATION NUMBER: 39,048
 REFERENCE/DOCKET NUMBER: 2418/9
 TELECOMMUNICATION INFORMATION:

Query Match 80.0%; Score 3222.5; DB: 4; Length 774;
 Best Local Similarity 77.4%; Pred. No. 6.5e-243;
 Matches 598; Conservative 83; Mismatches 89; Indels 3; Gaps 2;
 Db 1 MIFDMDYTRKDGPOLIRTFKENCEGKFTELDPHQPYYALLKDSATEIKAKGERH 60
 Qy 61 KIVRVWDAVKKKFLGRUVEVNLIFEPQDPALKQKIREPAVTDIYEKDFPEAKR 120
 Db 61 TUVVTKRVEVKQKKFLGRUVEVNLKLYFTHQDPAIRDKIREHGAVIDYEVDFPEAKR 120
 Qy 121 LIDKGLIPMEGDEELKIMAFDIEPYHGGDFEKGSEKGEIMISYADEEEARVITWNKIDPY 180
 Db 121 LIDKGLIPMEGDEELKIMAFDIEPYHGGDFEKGSEKGEIMISYADEEEARVITWNKIDPY 180
 Qy 181 VDVVSNERREMIKREVOIRKEKPDVLITYNGDNDFDPLYLIKRAEKGVLILLGRDKRPE 240
 Db 181 VDVVSTERMIKREPLRVKFDPLVITYNGDNDFDPLYLIKRAEKGVLNFAAGRDGS--E 238
 Qy 241 PTKHRMGDSFAVETKGRHIDFLFVVRRTINLPHYTLEAVYAVLGKTRSKLGAETAAI 300
 Db 239 PKIORMGDRAFEVKGRHIDFLPVIRTFINLPHYTLEAVYAVFGQPKKEVAAETPPA 298
 Qy 301 WETERSMKLAQSMEDAKVATYELGKEFLPMEAOQLSLIGOSLWDSRSSTGMLNVEWLL 360
 Db 299 WETGENLERVARYSMEDAKVATYELGKEFLPMEAOQLSLIGOSLWDSRSSTGMLNVEWLL 358
 Qy 361 RVAYERNELAPNPKDEEEYRRRLTYLGYWPERGLWENTYLDRCLYSLIVHN 420
 Db 359 RKAYERNELAPNPKDEKELARR-ROSVEGGYVPERGLWENTYLDRSLYSLIITHN 417
 Qy 421 VSPDTLERTGCKKNVAPTYGVYKCFDPGFPIISLGEILTMQOEIKKKMATIDPIEK 480
 Db 418 VSPDTLNREGCKEVDAPYGVHFRCKDPGFISLGLDLEERQIKKKMATIDPIEK 477
 Qy 481 MLDYRQRAVKLHANSYGYGMGPKARWYSKECSESVTAWGRHYTEMTKEIEKFGKV 540
 Db 478 LLDTRQRAKILANSYGYGPKARWYKCECAESVTANGREYITMTKEIEKFGKV 537
 Qy 541 YADDGFYATIPGKRPETIKKKAKEFLYINSKPLGULEYEGFYLGFVAKRYVI 600
 Db 538 YSDTDGFFATIPGADAETVKKAMFLNYINAKLPGALEYEGFYLGFVAKRYVI 597
 Qy 601 DEGRTTGLEVVRDWESEIAKETOAKYLEAIKEDSVAKEVAKIVDVEETAKYQPL 660
 Db 598 DEGKITTGLEVVRDWESEIAKETOARVLEALLKDGVEKAVRKVTEKLSKYEVPP 657
 Qy 661 EKLYTHEQITKDKDLYKATGPHVAKRLLAAG3KVRPOTISIVLRCGSKSDRVL 720
 Db 658 EKLYTHEQITRDLKDYKATGPHVAKRLLAAG3KVRPOTISIVLRCGSKSDRVL 717
 Qy 721 SEYDPKKKHYPDYYIENQVPAVLRLAEGYRKEDLYKQSSKQVGLDAWLK 773
 Db 718 DEFDPTKHKYDAEYIENQVPAVERILRAGFIRKEDLYKQTKRQVGLSAWLK 770

RESULT 7
 US-08-906-925-4
 ; Sequence 4, Application US/08906925
 ; Patent No. 5882904

GENERAL INFORMATION:

APPLICANT: Riedl, William

APPLICANT: Fly, Susan

APPLICANT: Kabord, Barbara F.

APPLICANT: Nye, Steven H.

APPLICANT: Ting, Eve

TITLE OF INVENTION: THERMOCOCCUS BAROSII DNA POLYMERASE MUTANTS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSSEE: Quarles & Brady

STREET: 411 East Wisconsin Avenue

CITY: Milwaukee

STATE: Wisconsin

NAME: Baker, Jean C.

REGISTRATION NUMBER: 35,433

REFERENCE/DOCKET NUMBER: 740211.90628

TELEPHONE: (414) 277-5709

TELEFAX: (414) 271-3552

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 778 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

S-08-906-925-4

ATTORNEY/AGENT INFORMATION:

NAME: Baker, Jean C.

REGISTRATION NUMBER: 35,433

REFERENCE/DOCKET NUMBER: 740211.90628

TELEPHONE: (414) 277-5709

TELEFAX: (414) 271-3552

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 778 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

S-08-906-925-4

Query Match 79.7% Score 3209.5 DB 2; Length 778; Best Local Similarity 77.2%; Pred. No. 6.8e-22; Mismatches 596; Conservative Matches 596; Indels 3; Gaps 2;

Y - 1 MFDPTDVTKGSKPIRIFKNGEKFKEILDPHFQPYIYALKDSAIDEITKAIGSKERHG 60

Y - 1 MLDVYDTEODCKPVIRFKKGEFKIEVDEPFPYIYALKDSAIDEITKAIGSKERHG 60

Y - 61 KIVRVVDAVKVKFLGRDVEWKFLFEPQDPAKGKIRHPAVDIDYEDIPFAKRY 120

Y - 61 KIVRVVDAVKVKFLGRDVEWKFLFEPQDPAKGKIRHPAVDIDYEDIPFAKRY 120

Y - 121 LDKGGLPMEGDEELKIMAFDTEFHEDGDERGKGEIMTSYADEEPAWTWKNDLPY 180

Y - 121 LDKGGLPMEGDEELKIMAFDTEFHEDGDERGKGEIMTSYADEEPAWTWKNDLPY 180

Y - 181 VDWSNREMIRKPVQVREKPDVLYINGNDFLPLIKRAEKUGVTLIGRKEHPE 240

Y - 181 VDWSNREMIRKPVQVREKPDVLYINGNDFLPLIKRAEKUGVTLIGRKEHPE 240

Y - 301 WTEESMKKLYASMEDARATELGKEFPMEAEALKLIGQSSWYDVSRSSTGNGNLVWYLL 360

Y - 301 WTEESMKKLYASMEDARATELGKEFPMEAEALKLIGQSSWYDVSRSSTGNGNLVWYLL 360

Y - 241 PKTHRMGDSFAVEBKGRIHDFLFPVVRTINPTYLEAVVAVLGTKSKGAEETAI 300

Y - 239 PKTQRMGDRFAVEVKGRHDFLIPVVRTINPLTYLEAVVAVLFGKPKWYAEIATA 298

Y - 299 RAYERMLNPKDRELAR-RGGYAGGYKEPERGLWNLWVYDLSRFSIYPSIITHN 417

RESULT 8

US-07-966-278-1

; Sequence 1, Application US/07966278

; Patent No. 549523

; GENERAL INFORMATION:

; APPLICANT: Mathur, Eric A

; TITLE OF INVENTION: EXONUCLEASE-DEFICIENT THERMOSPILATE

; TITLE OF INVENTION: PROCOCCUS FURIOSUS DNA POLYMERASE I

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSSEE: Thomas Fitting

; STREET: 12526 High Bluff Road, Suite 300

; CITY: San Diego

; STATE: CA

; COUNTRY: USA

; ZIP: 92130

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; APPLICATION NUMBER: US/07/966,278

; FILING DATE: 1992/2/26

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Fitting, Thomas

; REGISTRATION NUMBER: 34,163

; REFERENCE/DOCKET NUMBER: STC0133P

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619-792-3680

; TELEFAX: 619-792-8477

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 775 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; US-07-966-278-1

Query Match 78.3% Score 3151; DB 1; Length 775;

Matches 576; Conservative 100; Mismatches 94; Indels 2; Gaps 1;

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn, version #1.0, version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/424,921
FILING DATE: 19-APR-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,073
FILING DATE: 19-FEB-1991

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,568
FILING DATE: 03-DEC-1990

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/776,552
FILING DATE: 15-OCT-1991

ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas

REFERENCE/DOCKET NUMBER: STG0100P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-792-3680
TELEFAX: 619-792-8477

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 775 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTISENSE: NO

US-08-424-921-1

RESULT 9

US-08-424-921-1

Sequence 1, Application US/08424921

Patent No. 554552

GENERAL INFORMATION:

APPLICANT: Mathur, Eric A

TITLE OF INVENTION: PURIFIED THERMOSTABLE PYROCOCCUS
TITLE OF INVENTION: FURIOUS DNA POLYMERASE I

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Bingham & Fitting
STREET: 1256 High Bluff Road, Suite 300
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92130

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

Query Match 78.3%; Score 3151; DB 1; Length 775;
Best Local Similarity 74.6%; Pid. No. 2.4e-237;
Matches 576; Conservative 100; Mismatches 94; Indels 2; Gaps 1;

QY 1 MIFDTDYITKDGPKIIRFFKENGKFEKTFELDPHQPIYALLKDSATDEIKAKGERHG 60
Db 1 MILVDYITTEEGKPKVIRLKKENGKFKLHDRTFRPYIYALLRDKSKTEEVKKTGERHG 60

QY 541 YADTDGYATIPGKPKETKKAKAFKLKVINSKIPGLIELEYSGPYLROFFVAKRYAVI 600
Db 539 YDGDGLVATIPGKPKETEKKAKAFVKVINSKIPGLIELEYSGPYKRGFFVKKRYAVI 598

QY 601 DEEGITTRGLEYVWRDWEIAKETQAKYLEAETKEDSEKAVBIVKDVVEEAKYQPL 600
Db 599 DEEGKVITRGLTIVRDWSEIAKETQARVLETILKHDVVEAARIVKEVIOKLANEIPP 658

QY 661 EKVYTHEQTKDSEYKAIGPHVIAKRAAKGKVKRGCTIISTIVLGSKISDRVILL 720
Db 659 EKLAIEQITRPLHYKAIGPHVAVAKLAAGKVIKPOMVIGIVLVRDGPISNRATLA 718

QY 721 SEKPKKKHYPDPYIENQVLPVLRLAFGRKEDIKYQSSQVGLDAW 772
Db 719 EEXDQKKHXYDAEVTENQVLPVLRLAEGFVTKRQVGLTSWL 770

QY 181 VDVVSNERMIKREVOVREKDPDITYNGDSDFPVAKRABKLGKLTGRDGs-E 238

Db 181 VEVVSSERMIKFLRIRIETREKDPDITYNGDSDFPVAKRABKLGKLTGRDGs-E 238

QY 241 PKHRMGDSFAVEKGRHFDLFPVVRPITNLPIYTLAEVYAVLGKVKSLGAEETAI 300
Db 249 PKMORIGDMTAVEVKGRHFDLFPVTRITNLPIYTLAEVYALFGKPKKVVADELAKA 298

QY 301 WETEESMKLQSMSMEDAKATYELGKFPMEBLAKLGQSTWDVSSTSGNIVEWFL 360
Db 299 WESGENLERVAKYTSMEDAKATYELGKFPMEBLAKLGQSTWDVSSTSGNIVEWFL 358

QY 361 RVAYERNEALPNKDEEERRRTYLGKFPMEBLAKLGQSTWDVSSTSGNIVEWFL 420
Db 359 RVAYERNEALPNKDEEERRRTYLGKFPMEBLAKLGQSTWDVSSTSGNIVEWFL 418

QY 421 VSPDTLERGCKNQDVAPIVGYKCKDPGPFISLGLTILGELITMROEIKKMATDPIEK 480
Db 419 VSPDTLNEGCKNQDVAPIVGYKCKDPGPFISLGLHLEEROKIKTMKETDPIEK 478

QY 481 MLDRQRQAVKLHANSYYGMPKARWYSKECAESVTAWGRHYIEMTKEIEKFGFV 540

TOPLOGY: linear
; MOLECULE TYPE: protein
; US-08-556-355A-1

Query Match 78.3%; Score 3151; DB 2; Length 775;
Best Local Similarity 74.6%; Pred. No. 2, 4e-237; Mismatches 94; Indels 2; Gaps 1
Matches 576; Conservative 100; Mismatches 94; Indels 2; Gaps 1

Db 541 YADTGGYATPGKEPKTIKKAKEFKYINSKLPGLELEVEGYFLRGFVAKKYYAV 600
Db 539 YIDTGLYATPGESBEIKKAKFVKYINSKLPGLELEVEGYFKRGFFVTKKRYAVI 598

Db 601 DEEGRITRGLEVVRWRSLEKETAKVLEILKDSVRAVEIWKDVNEIAKQVL 660
Db 599 DEBGGVITRGLBVRDSEFAKETQARVLETFLKGDVEAVRIVKETOKLANEIPP 658

Db 661 EKLVHROITKOLSEYKAIGPHVAKRLLAANGKIKYRPGTTISYVLRGSKISDRVLL 720
Db 659 EKLAIEQITRPLHEKAIGPHVAVAKKLAANGVKIKPGMIVGIVLRGDOPISNAILA 718

Db 721 SEYDPKKHKKYDPYYENQVLPAVRLAEGYRKDLKQSSKGQGLDAML 772
Db 719 EYDPKKHKKYDAEYYENQVLPAVRLLEGFGYRKEDLRYOKTROVGLTSWL 770

RESULT 10
US-08-556-355A-1
Sequence 1, Application US/08556355A
Patent No. 5866395
GENERAL INFORMATION:
APPLICANT: MATHUR, Eric J.
TITLE OF INVENTION: Purified Thermostable Pyrococcus
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 08/08/556, 355A
FILING DATE: 13-NOV-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/424, 921
FILING DATE: 19-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/803, 627
FILING DATE: 02-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779, 846
FILING DATE: 21-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657, 073
FILING DATE: 19-FEB-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620, 568
FILING DATE: 03-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Barker, M. Paul
REGISTRATION NUMBER: 332, 013
REFERENCE/DOCKET NUMBER: 04121.0004-02
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)408-4400
TELEFAX: (202)408-4400
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 775 amino acids
TYPE: amino acid
STRANDEDNESS: single

Db 61 KIVRIVDVAVKKKFELDKRQVPAVRLAEGYRKDLKQSSKGQGLDAML 720
Db 61 KIVRIVDVAVKKKFELDKRQVPAVRLAEGYRKDLKQSSKGQGLDAML 718

Db 121 LTDKGJLPMEGDBEKLKMAFDTEFHEGDEFGKGKETIMISYADEEARVITWNKNDLY 180
Db 121 LTDKGJLPMEGDBEKLKMAFDTEFHEGDEFGKGKETIMISYADEEARVITWNKNDLY 180

Db 181 VDVVSNEREMTRKFVQVVRKEDPDLITYNGDNFDIPLYLIKRAEKAGVTLGDRKEHPE 240
Db 181 VEVVSNEREMTRKFVQVVRKEDPDLITYNGDNFDIPLYLIKRAEKAGVTLGDRKEHPE 238

Db 241 PKHMGDSFAVEIKRHFQFLPVVTRNPTYLEAVAEFLGKPLTKSGLAESTAI 300
Db 239 PRMQRIGDMTAVEVKSRHDFLYHVTTRNPTYLEAVAEFLGKPLTKSGLAESTAI 298

Db 301 WTEEESMKLQAQSYMDPARDATVELGKBFPMELAKLIGQSVDWSRSSTGNLVWYLL 360
Db 299 WESGENLNERVAKTSMEDPAKAYELGKBFPMELQSLRGQPLWDYRSSTGNLVWYLL 358

Db 361 RYAYERMELAPKPKDPEYRRLRTTYLGCVYKEPGLWNIITYDFRCYLPSTIHN 420
Db 359 RKAYERNEVAPNPKPSPEEYQRLRLESYTGCVYKEPGLWNIIVDFRCYLPSTIHN 418

Db 421 VSPDTLREGGNYDPAVQKFCDFGPGTIPSIGELIMPROEKKMATICDIEK 480
Db 419 VSPDTLNELEGCKNYDPAVQKFCDFGPGTIPSIGELIMPROEKKMATICDIEK 478

Db 481 MLYDRORAKVLHANSYYGGMGPKARWYKSCAESVTAWGRHYIEMTIKEEKEFKVL 540
Db 479 LLDYRORAKVLHANSYYGGMGPKARWYKSCAESVTAWGRHYIEMTIKEEKEFKVL 538

Db 541 YADTGGYATPGKEPKTIKKAKEFKYINSKLPGLELEVEGYFLRGFVAKKYYAV 600
Db 539 YIDTGLYATPGESBEIKKAKFVKYINSKLPGLELEVEGYFKRGFFVTKKRYAVI 598

Db 601 DEEGRITRGLEVVRWRSLEKETAKVLEILKDSVRAVEIWKDVNEIAKQVL 660
Db 599 DEBGGVITRGLBVRDSEFAKETQARVLETFLKGDVEAVRIVKETOKLANEIPP 658

Db 661 EKLVHROITKOLSEYKAIGPHVAKRLLAANGKIKYRPGTTISYVLRGSKISDRVLL 720
Db 659 EKLAIEQITRPLHEKAIGPHVAVAKKLAANGVKIKPGMIVGIVLRGDOPISNAILA 718

Db 721 SEYDPKKHKKYDPYYENQVLPAVRLAEGYRKDLKQSSKGQGLDAML 772
Db 719 EYDPKKHKKYDAEYYENQVLPAVRLLEGFGYRKEDLRYOKTROVGLTSWL 770

RESULT 11
US-07-803-627A-1
Sequence 1, Application US/07803627A
Patent No. 598663
GENERAL INFORMATION:
APPLICANT: MATHUR, Eric J.
TITLE OF INVENTION: Purified Thermostable Pyrococcus
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

ADDRESSEE: Dunner, L.L.P.
 STREET: 1300 I Street, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005

COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/803, 627A
 FILING DATE: 02-DEC-1991
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/776, 552
 FILING DATE: 14-OCT-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/657, 073
 FILING DATE: 19-FEB-1991

ATTORNEY/AGENT INFORMATION:
 NAME: Barker, M. Paul
 REGISTRATION NUMBER: 32, 013
 REFERENCE/DOCKET NUMBER: 04121. 0004-00
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)408-4000
 TELEFAX: (202)408-4400

INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 775 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-803-627A-1

Query Match 78.3%; Score 3151; DB 2;
 Best Local Similarity 74.6%; Pred. No. 2.4e-237; Length 775;
 Matches 576; Conservative 100; Mismatches 94; Indels 2; Gaps 1;

Y 1 MIRDFTDYITKDGPIIRIFKKGKNGEKFELDPHQPYIYALIKDDSADEIKAIIKGHRG 60
 Y 1 MIDDVDTITEEGPKVIRKFKKEGKFKTEHDRFRPPYIYALIKDDSKTEEVKKITGERH 60

QY 61 KIVVVDVAVKKFKFLGRDVEVKKLIFEPHQDPALRKIREHPAVDIDYEDVIFAKRY 120
 QY 61 KIVVIVDEVKEVKFLGRKPTVWKLYLEHPQDVPTIREVREHPAVWDIFEDVIFAKRY 120

QY 121 LIDKGKIPMEGDBELKLMADFETFYHGDEFCKGKGEIMISADEEEARVIVKNDIPLY 180

Db 121 LIDKGKIPMEGDBELKLMADFETFYHGDEFCKGKPIIMISADEEEARVIVKNDIPLY 180

Db 181 VEVVSSEREMIKFLRIRKEKDIDIVYNGDFDFPLAKRKEKGKLTGRGS-E 238

QY 241 PKIHRMGSFAVIIKGRIHFDPFLPVVRTINPFTYLVAYEAVLGKPKSKLGAEETAI 300
 QY 239 PKMQRIGDMTAVLVKGRIHFDPFLHVITRINPFTYLVAYEAVLGKPKKEVYADEIKA 298

QY 301 WEEFESMKKLAQYSMEDARATELGKEFPMEELAKLIGOSWVDSRSSTGNLVEWLL 360
 QY 299 WISSEGENLERYKSMEDAKATELGKEFPMEELAKLIGOSWVDSRSSTGNLVEWLL 358

QY 361 RYAYERNELAPNHPDEEYRRRTYLGIVPERGLWENYIYDPLVDSRSSTGNLVEWLL 420
 Db 359 RYAYERNELAPNHPSEEFQRRIRESYTGFFWKEPEKLWENYIYDPLVDSRSSTGNLVEWLL 418

QY 421 VSPDTLERSCKKNVDAPVGYKCFCKDFGFFIPSILGETITMREBKKMKTIDPIEK 480
 QY |||||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 419 VSPDTLNSGCKNVDAPQVGHCKDGFIPSULGHLEEROKKIKMKETQDPIEW 478
 QY 481 MLDYORAVKLHANSYGGNGYKARWYKRECASEVTANGRYHTEMTRKEIEFGPKV 540
 Db 479 LLDTROKAKLLANSFGYGYAKARWICKECAESVTANGRYKTELWKELEERFGPKV 538

QY 541 YADPDGFYITPGKPKETKKAKELFKINSKLPGLLEVEYFLRQFVAKRYAV 600
 Db 599 DEBKSVITNGLEIVRRDWBAIKEQARVLETLKHGDYFEEAVRIVKEVIOKLANEIP 658

QY 661 EKLYTHEQITKDSISBYKATGPHVATAKRAKAGTKVRCGTTISIVLRCGSKTSDRVIL 720
 Db 659 EKLAIVEQITRPLHEYKAGPHVAVAKKLAAGKIKPKMVIGIVLRCGSKTSDRVIL 718

QY 721 SEYPDKKKHHDYDPDYYENQVPAVLRLAFGYKREDLTYOSSKQVGLAHL 772
 QY |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||:|||:
 Db 719 EYDPKHHYDAEYIENQVPAVLRLAFGYKREDLTYOSSKQVGLAHL 770

RESULT 12

US-08-688-649-37

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

</div

Query Match 78.3%; Score 3151; DB 2; Length 776;
Best Local Similarity 74.6%; Pred. No. 2,37; Length 776;
Matches 576; Conservative 101; Mismatches 94; Indels 2; Gaps 1;

Qy 1 MIFDTDYITKGKPITRIFKENGKPFKIELDPHQPYVALLKDDSAIEIKAIKGHRG 60
Db 1 MILDVYITBEGKPVTRFLKKENKFKIEHDTFRPYTALLRDSKIEBVKKITGERH 60

Qy 61 KIVRVDAVKKKFLGRDVEVWKLIFEHQDVLALRKTRREHAVIDYEYDIPFAKR 120
Db 61 KIVRIVDVEREKFLGPITWKLYLEHQDVPITREKVREHPAVDIFEVYDIPFAKR 120

Qy 121 LIDKGKLIPMGDEEIKLMDIETFYHEDEFGKEITMISYABEEARVITWKNDLPY 180
Db 121 LIDKGKLIPMGEEIKLMDIETFYHEDEFGKEITMISYABEEARVITWKNDLPY 180

Qy 181 VDVVSNEREMIKRFQVIVREKDPDVLTNGDNFDPLYLUKRAEKGVLGVLGKHEPE 240
Db 181 VEVVSNEREMIKRFQVIVREKDPDVLTNGDNFDPLYLUKRAEKGVLGKLITIGRGS-E 238

Qy 241 PKIHRMGDSFAVEKGRHIDLFPVVRTNLPTYLEAVYAVIGKTISKLGEEIAI 300
Db 239 PKMQRIGDWTAVEWGRHIDLYHVTINTLPYLTLEAVEYAFGKPREKVADIEKA 293

Qy 301 WETESMKLAQYSMEDARATYELGKEFMEALAKLIGQSVMDVSRSSTGNLVEWLL 360
Db 299 WESGENLERYAKYSPMEDAKATYELGKEFMEALAKLIGQSVMDVSRSSTGNLVEWLL 358

Qy 361 RVAYERNELAPNPKPDEEVRRRLRTYLGKVKEPERGLMENTYLDFLRCLYPSIVTHN 420
Db 359 RKAVERNEVAPNPKPDEEVRRRLRTYLGKVKEPERGLMENTYLDFLRCLYPSIVTHN 418

Qy 421 VSPDTLREGECKKNVQVAPTYGKCKDFPOFIPSILGELITMRQBIKKMKTDPIEK 480
Db 419 VSPDTLREGECKKNVQVAPTYGKCKDFPOFIPSILGELITMRQBIKKMKTDPIEK 478

Qy 481 MLDYQRAVKHLANSYGGYGMKPCKARWYCKECAESVTANGRKYLUWKLLEERKGFKV 540
Db 479 LLDYQRAKTLANSFYGYYGKARWYCKECAESVTANGRKYLUWKLLEERKGFKV 538

Qy 541 YADTQGYATIPGEKPKETKKAKFLYTKNSKPGGLELEYEGYFLRGFVAKKRYV 600
Db 539 YIDTQGYATIPGEKPKETKKAKFLYTKNSKPGGLELEYEGYFLRGFVAKKRYV 598

Qy 601 DEEGRITTRGLEYVWDRDSETAKEQVAKLLEALKEDSVKAVEVKDVVEELAKYQPL 660
Db 599 DEEGKVTGLEYVWDRDSETAKEQVAKLLEALKEDSVKAVEVKDVVEELAKYQPL 658

Qy 661 EKLVINEHQITKDLSEKATGPHVAKRAKKGIVRPGILISYVLRGSKISRVILL 720
Db 659 EKLVATEQVITRPLHRYKATGPHVAKKLAKGKVKPGHVGIVYTLRGDPINRAILA 718

Qy 721 SEYDPKKHYPDQYTTENOYLPAVIRILEAFGYRVEDLKVQSSKGVLDAWL 772
Db 719 BEYDQKKHYPDQYTTENOYLPAVIRILEAFGYRVEDLKVQSSKGVLDAWL 770

RESULT 13
US-08-375-134-12
; Sequence 12, Application US/08375134
; Patent No. 5,602011
GENERAL INFORMATION:
APPLICANT: Lubin, Robert A.
APPLICANT: d'Hennzel, Olga B.
APPLICANT: Duffaud, Guy D.
APPLICANT: Jolly, James F.
APPLICANT: Kelly, Robert M.
APPLICANT: Ting, Eve Yl-Fay
TITLE OF INVENTION: PURIFIED THERMOCOCCUS BAROSII DNAPOLYMERASE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 411 East Wisconsin Avenue

CITY: Milwaukee
STATE: Wisconsin
COUNTRY: U.S.A.
ZIP: 53202-497

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-POS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/375,134
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Baker, Jean C.
REGISTRATION NUMBER: 35,433
REFERENCE/DOCKET NUMBER: 740211-90440
TELECOMMUNICATION INFORMATION:
TELEPHONE: (414) 277-5709
TELEFAX: (414) 271-3552
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 779 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-375-134-12

Query Match 78.0%; Score 3139; DB 1; Length 779;
Best Local Similarity 76.2%; Pred. No. 2,1e-236; Length 779;
Matches 589; Conservative 76; Mismatches 99; Indels 4; Gaps 3;

Qy 1 MIFDTDYITKGKPITRIFKENGKPFKIELDPHQPYVALLKDDSAIEIKAIKGHRG 60
Db 1 MILDVYITBEGKPVTRFLKKENKFKIEHDTFRPYTALLRDSKIEBVKKITGERH 60

Qy 61 KIVRVDAVKKKFLGRDVEVWKLIFEHQDVLALRKTRREHAVIDYEYDIPFAKR 120
Db 61 KIVRIVDVEREKFLGPITWKLYLEHQDVPITREKVREHPAVDIFEVYDIPFAKR 120

Qy 121 LIDKGKLIPMGDEEIKLMDIETFYHEDEFGKEITMISYABEEARVITWKNDLPY 180
Db 121 LIDKGKLIPMGDEEIKLMDIETFYHEDEFGKEITMISYABEEARVITWKNDLPY 180

Qy 181 VDVVSNEREMIKRFQVIVREKDPDVLTNGDNFDPLYLUKRAEKGVLGVLGKHEPE 240
Db 181 VEVVSNEREMIKRFQVIVREKDPDVLTNGDNFDPLYLUKRAEKGVLGKLITIGRGS-E 238

Qy 241 PKIHRMGDSFAVEKGRHIDLFPVVRTNLPTYLEAVYAVIGKTISKLGEEIAI 300
Db 239 PKMQRIGDWTAVEWGRHIDLYHVTINTLPYLTLEAVEYAFGKPREKVADIEKA 293

Qy 301 WETESMKLAQYSMEDARATYELGKEFMEALAKLIGQSVMDVSRSSTGNLVEWLL 360
Db 299 WESGENLERYAKYSPMEDAKATYELGKEFMEALAKLIGQSVMDVSRSSTGNLVEWLL 358

Qy 361 RVAYERNELAPNPKPDEEVRRRLRTYLGKVKEPERGLMENTYLDFLRCLYPSIVTHN 420
Db 359 RKAVERNEVAPNPKPDEEVRRRLRTYLGKVKEPERGLMENTYLDFLRCLYPSIVTHN 418

Qy 421 VSPDTLREGECKKNVQVAPTYGKCKDFPOFIPSILGELITMRQBIKKMKTDPIEK 480
Db 419 VSPDTLREGECKKNVQVAPTYGKCKDFPOFIPSILGELITMRQBIKKMKTDPIEK 478

Qy 481 MLDYQRAVKHLANSYGGYGMKPCKARWYCKECAESVTANGRKYLUWKLLEERKGFKV 540
Db 479 LLDYQRAKTLANSFYGYYGKARWYCKECAESVTANGRKYLUWKLLEERKGFKV 538

Qy 541 YADTQGYATIPGEKPKETKKAKFLYTKNSKPGGLELEYEGYFLRGFVAKKRYV 600
Db 539 YIDTQGYATIPGEKPKETKKAKFLYTKNSKPGGLELEYEGYFLRGFVAKKRYV 598

Qy 601 DEEGRITTRGLEYVWDRDSETAKEQVAKLLEALKEDSVKAVEVKDVVEELAKYQPL 660
Db 599 DEEGKVTGLEYVWDRDSETAKEQVAKLLEALKEDSVKAVEVKDVVEELAKYQPL 658

Qy 661 EKLVINEHQITKDLSEKATGPHVAKRAKKGIVRPGILISYVLRGSKISRVILL 720
Db 659 EKLVATEQVITRPLHRYKATGPHVAKKLAKGKVKPGHVGIVYTLRGDPINRAILA 718

Qy 721 SEYDPKKHYPDQYTTENOYLPAVIRILEAFGYRVEDLKVQSSKGVLDAWL 772
Db 719 BEYDQKKHYPDQYTTENOYLPAVIRILEAFGYRVEDLKVQSSKGVLDAWL 770

Qy 300 WETESMKLAQYSMEDARATYELGKEFMEALAKLIGQSVMDVSRSSTGNLVEWLL 359
Db 299 AWETGEGLEGYARYSMEDARATYELGKEFMEALAKLIGQSVMDVSRSSTGNLVEWLL 358

Qy 360 RVAYERNELAPNPKPDEEVRRRLRTYLGKVKEPERGLMENTYLDFLRCLYPSIVTHN 420
Db 359 RKAVERNEVAPNPKPDEEVRRRLRTYLGKVKEPERGLMENTYLDFLRCLYPSIVTHN 417

Qy 420 VSPDTLREGECKKNVQVAPTYGKCKDFPOFIPSILGELITMRQBIKKMKTDPIEK 479
Db 419 VSPDTLREGECKKNVQVAPTYGKCKDFPOFIPSILGELITMRQBIKKMKTDPIEK 477

Qy 480 KLDYQRAVKHLANSYGGYGMKPCKARWYCKECAESVTANGRKYLUWKLLEERKGFKV 540
Db 478 KLDYQRAKTLANSFYGYYGKARWYCKECAESVTANGRKYLUWKLLEERKGFKV 537

Qy 540 LYADTQGYATIPGEKPKETKKAKFLYTKNSKPGGLELEYEGYFLRGFVAKKRYV 599
Db 538 LYADTQGYATIPGEKPKETKKAKFLYTKNSKPGGLELEYEGYFLRGFVAKKRYV 597

APPLICATION NUMBER: US 07/686,340
 FILING DATE: 17-ARR-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/626,057
 FILING DATE: 11-DEC-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/513,994
 FILING DATE: 26-APR-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: WILLIAMS, GREGORY D.
 REGISTRATION NUMBER: 30901
 REFERENCE/DOCKET NUMBER: NEB-101
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (508) 927-5054
 TELEFAX: (508) 927-1705
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1022 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: protein

Query Match 62.2%; Score 2502.5; DB 1; Length 1022;
 Best Local Similarity 62.9%; Pred. No. 1.2e-186; Indels 127; Gaps 17;
 Matches 534; Conservative 59; Mismatches 129;

Qy 1 MIFDPTYITKQKPIRIFKENGKESKIELDPHQQYIYALLKDDSAIDEKAIKGERRHG 60
 Db 1 MILDPTYITKQKPIRIFKENGKESKIELDPHQQYIYALLKDDSAIDEKAIKGERRHG 60

Qy 61 KIVRVVDAVKKKFGLGRDVEWKWLIFEPDPAVLRKGKREHAPAVDIEYDIFPAKRY 120
 Db 61 KTVRVVDAVKYRKFLGREGEWVKLIFEPDPAVLRKGKREHAPAVDIEYDIFPAKRY 120

Qy 121 LIDKGKLIPMEEDEELKLMADFETEVHEGDFGKGETIMSYADEREARVITWKNTDLPY 180
 Db 121 LIDKGKLIPMEEDEELKLMADFETEVHEGDFGKGETIMSYADEREARVITWKNTDLPY 180

Qy 181 VDVVSVEREMKRFQYIVREKDPDVITYGDNFDLPLYLKRAEKYGVTLLGROKEHPE 240
 Db 181 VDVVSNEREMKRFQYIVREKDPDVITYGDNFDLPLYLKRAEKYGVTLLGROKEHPE 240

Qy 241 PKIRHGDFAVEKGRHIEDLPVWRRTINLPTYLEAVNVLQTKSKLGAEETAI 300
 Db 241 PKIRHGDFAVEKGRHIEDLPVWRRTINLPTYLEAVNVLQTKSKLGAEETAI 300

Qy 301 WETEESMKKLAQYSMEDARATELYGREGFPMEALKLIGOSWMDVSRSSIGNLYEWYL 360
 Db 301 WETEESMKKLAQYSMEDARATELYGREGFPMEALKLIGOSWMDVSRSSIGNLYEWYL 360

Qy 361 RVAYEERNEALANKPDEEYERRLRTYLGKVKKEPERGLMENTITYDFCLYPSLIVTHN 420
 Db 361 RVAYEERNEALANKPDEEYERRLRTYLGKVKKEPERGLMENTITYDFCLYPSLIVTHN 420

Qy 421 VSPDTLEREGCKNYDAPIVQYKFCDFGPGTIPSTIGELITMRQEKKKMATTIPIEK 480
 Db 421 VSPDTLEREGCKNYDAPIVQYKFCDFGPGTIPSTIGELITMRQEKKKMATTIPIEK 480

Qy 481 MLDYRORAKVLHANSY-----GYMGYPK----- 504
 Db 481 MLDYRORAKVLHANSY-----GYMGYPK----- 504

Qy 505 -----ARWYSKEAES---VTAWGHYIEMTIKEEKEFGFKV-----LYADPDGFY 548
 Db 541 VNNLAFAFSENKRKIKESEVVKVYKALIRHKYKRAVEIQLSSGRKINITAGHSLFVTNRGEI 600

Qy 549 ATIPG-----KPETIKKAKEFKYIINSKLPGLEEYEGYL-----RGFF- 591
 Db 601 KEVSGOGIKEGDLIVAPKKIKLNEKGVSINPELISLSDSEBETADIVMTISAKGRNFEK 660

Search completed: May 28, 2002, 14:17:33
 Job time: 31 sec

Qy 592 --VAKKRYAVIDEESGTT-----RGLEWVRDWSERAKETQKA-KVU 630
 Db 661 GMLRTLWMMFGBENRIRRTNRYLFLKEKLUKLILPGRYEV-----DWERLKKYKOLYKL 718
 Qy 631 EAIKEDSVRAVEVAKVQVPLERKLVHEQIKDSEYKAISPHVATKRLA 690
 Db 719 AGSVKNGNKRKBYLWNEKDFISY-FP-----OKELEMWK-ICFLNGFRNTNCI 766
 Qy 691 ARGIKVRPGTITSYIYLRG-----SGKTSDRVILLSEVDPKKHRYDPOYIENQVLP 742
 Db 767 LK-VDEDFGKULGYYVSEGVAQAKNKTGJS-----YSVLYNEDPN-----VLE 811

Qy 743 AVRILAEAF 751
 Db 812 SMKNVAEK 820